GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 14:57:06 ; Search time 35 Seconds
(without alignments)
1422.976 Million cell updates/sec

Title: US-09-813-775C-34
Perfect score: 991
Sequence: 1 MGVHECPAWLWLLLSLLSLP.....NFLRGKLKLYTGEACRTGDR 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	759.5	76.6	194	6	Q9MYM8	Q9mym8 sus scrofa
2	729	73.6	195	6	Q9GKA3	Q9gka3 oryctolagus
3	726.5	73.3	195	6	Q9GKA2	Q9gka2 oryctolagus
4	645	65.1	133	6	Q8HZ88	Q8hz88 gorilla gor
5	632	63.8	133	6	Q8HZ89	Q8hz89 pan troglod
6	601	60.6	131	6	Q8HZ87	Q8hz87 pongo pygma
7	581	58.6	133	6	Q8HZ86	Q8hz86 macaca sp.
8	533	53.8	133	6	Q8HZ85	Q8hz85 saguinus oe
9	188	19.0	50	11	Q9QV40	Q9qv40 rattus sp.
10	96.5	9.7	554	2	Q9RPH5	Q9rph5 mycobacteri
11	90	9.1	623	11	Q9EQR6	Q9eqr6 mus musculu
12	87.5	8.8	346	16	Q8Z2M5	Q8z2m5 salmonella
13	87.5	8.8	346	16	Q8ZKZ4	Q8zkz4 salmonella
14	87.5	8.8	454	11	Q9D7X0	Q9d7x0 mus musculu
15	87	8.8	871	5	Q9NB04	Q9nb04 drosophila
16	85	8.6	235	11	Q9QVR7	Q9qvr7 mus sp. tru
17	85	8.6	877	5	Q8T5Z7	Q8t5z7 dictyosteli
18	84	8.5	339	16	Q9HZM7	Q9hzm7 pseudomonas
19	84	8.5	378	17	Q8ZZK3	Q8zzk3 pyrobaculum
20	82.5	8.3	389	2	Q9AN18	Q9an18 bradyrhizob
21	81.5	8.2	436	3	Q8J1J9	Q8j1j9 saccharomyc
22	81.5	8.2	480	10	Q43380	Q43380 avena sativ
23	81	8.2	323	16	Q8ZDC8	Q8zdc8 yersinia pe
24	81	8.2	414	12	Q68818	Q68818 hepatitis c
25	80	8.1	522	4	Q8N3D3	Q8n3d3 homo sapien
26	80	8.1	655	5	Q94479	Q94479 dictyosteli
27	79.5	8.0	475	16	Q9K7P5	Q9k7p5 bacillus ha
28	79.5	8.0	896	2	Q9AN79	Q9an79 bradyrhizob
29	79.5	8.0	3722	2	P94873	P94873 lysobacter
30	79	8.0	194	10	Q8LHZ9	Q8lhz9 oryza sativ
31	79	8.0	743	16	Q9RYG5	Q9ryg5 deinococcus
32	78.5	7.9	296	16	Q8ZAY4	Q8zay4 yersinia pe
33	78.5	7.9	348	16	O86751	O86751 streptomyce
34	78.5	7.9	383	16	Q92QV7	Q92qv7 rhizobium m
35	78.5	7.9	2201	16	Q8ZHV3	Q8zhv3 yersinia pe
36	78	7.9	389	16	Q8UAR7	Q8uar7 agrobacteri
37	78	7.9	451	16	P74054	P74054 synechocyst
38	78	7.9	861	2	Q9RQS1	Q9rqs1 escherichia
39	78	7.9	1189	2	Q44415	Q44415 agrobacteri
40	77.5	7.8	242	16	Q8YY76	Q8yy76 anabaena sp
41	77.5	7.8	567	10	Q9SVL5	Q9sv15 arabidopsis
42	77.5	7.8	569	10	Q8LPQ3	Q8lpq3 arabidopsis
43	77.5	7.8	774	10	O64777	O64777 arabidopsis
44	77	7.8	416	17	Q9HQE1	Q9hqe1 halobacteri
45	77	7.8	613	5	Q9VJ39	Q9vj39 drosophila

ALIGNMENTS

RESULT 1
Q9MYM8
ID Q9MYM8 PRELIMINARY; PRT; 194 AA.
AC Q9MYM8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NORWEGIAN LANDRACE; TISSUE=Kidney;
RA David B., Harbitz I.;
RT "The porcine erythropoietin gene: cDNA and genomic sequences and
RT expression analyses.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ249745; CAB96416.1; -.
DR EMBL; AJ249746; CAB96417.1; -.
DR HSSP; P01588; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythroptn.
DR Pfam; PF00758; EPO_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO; 1.
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 194 ERYTHROPOIETIN.
SQ SEQUENCE 194 AA; 21303 MW; 77881A6F6F20EA1C CRC64;

Query Match 76.6%; Score 759.5; DB 6; Length 194;
Best Local Similarity 80.5%; Pred. No. 4.7e-65;
Matches 157; Conservative 7; Mismatches 28; Indels 3; Gaps 2;

Qy 1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

```
Db          1 MGARECPARL-LLLSLLLLPLGLPVLGAPPRLICDSRVLERYILEAKEGENATMGCAESC 59

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              | :|||||||||||||||||    |||:|||||||||||||:|:|||||  |||||  |  |||
Db         60 SFSENITVPDTKVNFYAWKRMEVQQQAMEVWQGLALLSEAILQGQALLANSSQPSEALQL 119

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDA--ASAAPLRTITADTFRKLFRVYSNFLRG 178
              ||||||||||||:||||||||||||  |||  :|| ||||   ||  |||| |||||||
Db        120 HVDKAVSGLRSLTSLLRALGAQKEAIPLPDASPSSATPLRTFAVDTLCKLFRNYSNFLRG 179

Qy        179 KLKLYTGEACRTGDR 193
              || ||||||||  ||
Db        180 KLTLYTGEACRRRDR 194


RESULT 2
Q9GKA3
ID   Q9GKA3      PRELIMINARY;      PRT;   195 AA.
AC   Q9GKA3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Erythropoietin.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21290682; PubMed=11396976;
RA   Vilalta A., Wu D., Margalith M., Hobart P.;
RT   "Rabbit EPO Gene and cDNA: Expression of Rabbit EPO after
RT   Intramuscular Injection of pDNA.";
RL   Biochem. Biophys. Res. Commun. 284:823-827(2001).
DR   EMBL; AF290943; AAG36961.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
SQ   SEQUENCE   195 AA;  21053 MW;  0999DA7D852713F3 CRC64;

  Query Match             73.6%;  Score 729;  DB 6;  Length 195;
  Best Local Similarity   77.9%;  Pred. No. 4.1e-62;
  Matches  152;  Conservative   14;  Mismatches   27;  Indels    2;  Gaps    2;

Qy          1 MGVHECPAWLWL-LLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEH 59
              |||    | | | || || | |||||||| ||||||||||||:||||||:| ||||
Db          1 MGVRGRLALLPLALLCLLVLALGLPVLGAPARLICDSRVLERYILEAKEAENVTMGCAEG 60

Qy         60 CSLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQ 119
              ||| ||||||||||||: ||:: : : |||||||||||||||:|| |||| ||||  | ||
Db         61 CSLGENITVPDTKVNFHHWKKSEAGRHAVEVWQGLALLSEAMLRSQALLANSSQLPETLQ 120

Qy        120 LHVDKAVSGLRSLTTLLRALGAQKEAISPPDAA-SAAPLRTITADTFRKLFRVYSNFLRG 178
              :|||||||||||||:|||||| |||||:|||:|| |||||||: |||  ||||:|||||||
Db        121 VHVDKAVSGLRSLTSLLRALGVQKEAVSPPEAASSAAPLRTVAADTLCKLFRIYSNFLRG 180

Qy        179 KLKLYTGEACRTGDR 193
              |||||||||| |||
Db        181 KLKLYTGEACRRGDR 195


RESULT 3
Q9GKA2
ID   Q9GKA2      PRELIMINARY;      PRT;   195 AA.
AC   Q9GKA2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Erythropoietin.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21290682; PubMed=11396976;
RA   Vilalta A., Wu D., Margalith M., Hobart P.;
RT   "Rabbit EPO Gene and cDNA: Expression of Rabbit EPO after
RT   Intramuscular Injection of pDNA.";
RL   Biochem. Biophys. Res. Commun. 284:823-827(2001).
DR   EMBL; AF290944; AAG36962.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
SQ   SEQUENCE   195 AA;  21025 MW;  1F1DC7F403A303EC CRC64;

  Query Match             73.3%;  Score 726.5;  DB 6;  Length 195;
  Best Local Similarity   80.2%;  Pred. No. 7.1e-62;
  Matches  146;  Conservative   14;  Mismatches   21;  Indels    1;  Gaps    1;

Qy         13 LLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTK 72
              || || | |||||||| ||||||||||||:||||||:| |||| ||| |||||||||
Db         14 LLCLLVLALGLPVLGAPARLICDSRVLERYILEAKEAENVTMGCAEGCSLGENITVPDTK 73

Qy         73 VNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSL 132
              |||: ||:: : : |||||||||||||:|| |||| ||||  | ||:|||||||||||
Db         74 VNFHHWKKSEAGRHAVEVWQGLALLSEAMLRSQALLANSSQLPETLQVHVDKAVSGLRSL 133

Qy        133 TTLLRALGAQKEAISPPDAA-SAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTG 191
              |:|||||| ||||:|||:|| |||||||: |||  ||||:|||||||||||||||||| |
Db        134 TSLLRALGVQKEAVSPPEAASSAAPLRTVAADTLCKLFRIYSNFLRGKLKLYTGEACRRG 193

Qy        192 DR 193
              ||
Db        194 DR 195


RESULT 4
Q8HZ88
ID   Q8HZ88      PRELIMINARY;      PRT;   133 AA.
AC   Q8HZ88;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Erythropoietin (Fragment).
OS   Gorilla gorilla (gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX   NCBI_TaxID=9593;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   O'hUigin C., Tichy H., Klein J.;
RT   "Molecular evolution in higher primates; gene specific and organism
RT   specific characteristics.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY092016; AAM76633.1; -.
FT   NON_TER       1      1
FT   NON_TER     133    133
SQ   SEQUENCE   133 AA;  14696 MW;  E2D0130942693140 CRC64;

  Query Match             65.1%;  Score 645;  DB 6;  Length 133;
  Best Local Similarity   97.0%;  Pred. No. 3e-54;
  Matches  129;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         36 SRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLA 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||
Db          1 SRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVRQQAVEVWQGLA 60

Qy         96 LLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAA 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 LLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAA 120

Qy        156 PLRTITADTFRKL 168
              |||||||||||||
Db        121 PLRTITADTFRKL 133


RESULT 5
Q8HZ89
ID   Q8HZ89      PRELIMINARY;      PRT;   133 AA.
AC   Q8HZ89;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Erythropoietin (Fragment).
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   O'hUigin C., Tichy H., Klein J.;
RT   "Molecular evolution in higher primates; gene specific and organism
RT   specific characteristics.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY092015; AAM76632.1; -.
FT   NON_TER       1       1
FT   NON_TER     133     133
SQ   SEQUENCE   133 AA;  14523 MW;  ECDF5609596390EB CRC64;

  Query Match             63.8%;  Score 632;  DB 6;  Length 133;
  Best Local Similarity   96.2%;  Pred. No. 5.3e-53;
  Matches  128;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         36 SRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLA 95
              |||||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||
Db          1 SRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLA 60

Qy         96 LLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAA 155
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         61 LLSEAVLRGQALLVNSSQPLEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAA 120

Qy        156 PLRTITADTFRKL 168
              |||||||||||||
Db        121 PLRTITADTFRKL 133


RESULT 6
Q8HZ87
ID   Q8HZ87      PRELIMINARY;      PRT;   131 AA.
AC   Q8HZ87;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Erythropoietin (Fragment).
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   O'hUigin C., Tichy H., Klein J.;
RT   "Molecular evolution in higher primates; gene specific and organism
RT   specific characteristics.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY092017; AAM76634.1; -.
FT   NON_TER       1       1
FT   NON_TER     131     131
SQ   SEQUENCE   131 AA;  14403 MW;  143F5E4931EA03FA CRC64;

  Query Match             60.6%;  Score 601;  DB 6;  Length 131;
  Best Local Similarity   92.4%;  Pred. No. 5e-50;
  Matches  122;  Conservative    2;  Mismatches    6;  Indels    2;  Gaps    1;

Qy         36 SRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLA 95
              |||||||||||||||  |||||||||:||||||||||||||||    |||:|||||||
Db          1 SRVLERYLLEAKEAEN--TGCAEHCSLSENITVPDTKVNFYAWKRMEVGQQALEVWQGLA 58

Qy         96 LLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAA 155
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||
Db         59 LLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEASSPPDAALAA 118

Qy        156 PLRTITADTFRK 167
              ||||||||||||
Db        119 PLRTITADTFRK 130


RESULT 7
Q8HZ86
ID   Q8HZ86      PRELIMINARY;      PRT;   133 AA.
AC   Q8HZ86;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Erythropoietin (Fragment).
OS   Macaca sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9549;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   O'hUigin C., Tichy H., Klein J.;
RT   "Molecular evolution in higher primates; gene specific and organism
RT   specific characteristics.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY092018; AAM76635.1; -.
FT   NON_TER       1       1
FT   NON_TER     133     133
SQ   SEQUENCE   133 AA;  14506 MW;  55CB11A2DC8354A0 CRC64;

  Query Match             58.6%;  Score 581;  DB 6;  Length 133;
  Best Local Similarity   87.2%;  Pred. No. 4.2e-48;
  Matches  116;  Conservative    8;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         36 SRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLA 95
              |||||||||||||||:| ||:| ||||||||||||||||||    ||||||||||
Db          1 SRVLERYLLEAKEAENVTMGCSESCSLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLA 60

Qy         96 LLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAA 155
              ||||||||||:| ||||:|||||:|||:|||||:||||||||||:|||| |||||||
Db         61 LLSEAVLRGQAVLANSSQPFEPLQLHMDKAISGLRSITTLLRALGAQQEAISLPDAASAA 120

Qy        156 PLRTITADTFRKL 168
              |||||||||| ||
Db        121 PLRTITADTFCKL 133


RESULT 8
Q8HZ85
ID   Q8HZ85      PRELIMINARY;      PRT;   133 AA.
AC   Q8HZ85;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Erythropoietin (Fragment).
OS   Saguinus oedipus (Cotton-top tamarin).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
OX   NCBI_TaxID=9490;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   O'hUigin C., Tichy H., Klein J.;
RT   "Molecular evolution in higher primates; gene specific and organism
RT   specific characteristics.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AY092019; AAM76636.1; -.
FT   NON_TER       1      1
FT   NON_TER     133    133
SQ   SEQUENCE   133 AA;  14375 MW;  C923E859BB608FEC CRC64;

  Query Match             53.8%;  Score 533;  DB 6;  Length 133;
  Best Local Similarity   82.1%;  Pred. No. 1.7e-43;
  Matches  110;  Conservative    8;  Mismatches   14;  Indels    2;  Gaps    2;

Qy         36 SRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLA 95
              | ||||||:|| ||||||:| |||| |||||||||||||||||||:    ||||:|||||
Db          1 SGVLERYVLEGKEAENVTMGCAESCSLNENITVPDTKVNFYAWKKMEFGQQAVDVWQGLT 60

Qy         96 LLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAA-SA 154
              |||||||||||||| ||:|| |||||||:|:|||||||||||||||||| || |||||| ||
Db         61 LLSEAVLRGQALLANSTQPREPLQLHMDRAVSGLRSLTTLLRALGAQ-EATSPPDAAPSA 119

Qy        155 APLRTITADTFRKL 168
               ||:|||||||| ||
Db        120 VPLQTITADTFSKL 133


RESULT 9
Q9QV40
ID   Q9QV40      PRELIMINARY;      PRT;    50 AA.
AC   Q9QV40;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Erythropoietin (Fragment).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94115047; PubMed=7764337;
RA   Okano M., Suga H., Masuda S., Nagao M., Narita H., Ikura K.,
RA   Sasaki R.;
RL   Biosci. Biotechnol. Biochem. 57:1882-1885(1993).
DR   HSSP; P01588; 1EER.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
SQ   SEQUENCE   50 AA;  5587 MW;  70B44A8BFE016034 CRC64;

  Query Match             19.0%;  Score 188;  DB 11;  Length 50;
  Best Local Similarity   78.0%;  Pred. No. 7.4e-11;
  Matches   39;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYA 77
              |||||| |||||||:|||||| :| | ||   |:| ||||||||||||
Db          1 APPRLIXDSRVLERYILEAKEAEXVTMGXAEGPRLSEXITVPDTKVNFYA 50


RESULT 10
Q9RPH5
ID   Q9RPH5      PRELIMINARY;      PRT;   554 AA.
AC   Q9RPH5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RecD.
GN   RECD.
OS   Mycobacterium smegmatis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1772;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=mc2155;
RX   MEDLINE=99412429; PubMed=10481025;
RA   Griffin IV T.J., Parsons L., Leschziner A.E., DeVost J.,
RA   Derbyshire K.M., Grindley N.D.;
RT   "In vitro transposition of tn552: a tool for DNA sequencing and
RT   mutagenesis.";
RL   Nucleic Acids Res. 27:3859-3865(1999).
DR   EMBL; AF157643; AAD46809.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR006344; RecD.
DR   InterPro; IPR000606; Viral_helicase1.
DR   Pfam; PF01443; Viral_helicase1; 1.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMs; TIGR01447; recD; 1.
KW   ATP-binding.
SQ   SEQUENCE   554 AA;  59516 MW;  3947B550498BD62C CRC64;

  Query Match              9.7%;  Score 96.5;  DB 2;  Length 554;
  Best Local Similarity   26.7%;  Pred. No. 0.9;
  Matches   50;  Conservative   19;  Mismatches   57;  Indels   61;  Gaps    9;

Qy          8 AWLWLLLSLLSLPLGLPVLGAPP--RLICDSRVLERYLLEAKEAENITTGCAEHCSL--- 62
              |||  | :        |:|| ||  ||  |   |:|| || ::       | :  :|
Db         72 AWLAALAA-------SPLLGQPPVLRLFGDLLYLDRYWLEEQQV------CDDVLALVSA 118

Qy         63 NENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSE-------------AVLRGQALL 108
                 |||    | |          | : ||| :|:               |:|  || |
Db        119 RPGGAVPDVSRLFGAGFEEQRAAAKVALSQGLTVLTGGPGTGKTTTVARLLALLAEQAAL 178

Qy        109 VNSSQP---------------WEPLQLHVD-------KAVSGLRSLTTLLRALGAQKEAI 146
                 |                | :|| :|       : ::|| : ||| | || :
Db        179 AGKPSPRIALAAPTGKAAARLQEAVQLEIDQLDLIERRRLTGLHA-TTLHRLLGPR---- 233

Qy        147 SPPDAAS 153
                 || :|
Db        234 --PDTSS 238


RESULT 11
Q9EQR6
ID   Q9EQR6      PRELIMINARY;      PRT;   623 AA.
AC   Q9EQR6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Fanconi anemia group G protein (Fanconi anemia complementation group
DE   G).
GN   FANCG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tipping A.J., Morgan N.V., Mathew C.G.;
RT   "Sequencing and genomic structure of fancg, the murine orthologue of
RT   the Fanconi anemia group G gene (FANCG).";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB;
RA   van de Vrugt H.J., Koomen M., Berns M.A.B., de Vries Y.,
RA   Rooimans M.A., van der Weel L., Blom E., de Winter J.P., de Groot J.,
RA   Schepers R.J., Hoatlin M.E., Ching Cheng N., Joenje H., Arwert F.;
RT   "Characterization, expression and complex formation of murine Fancg.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF112439; AAG43198.1; -.
DR   EMBL; AY049715; AAL12165.1; -.
DR   MGD; MGI:1926471; Fancg.
DR   InterPro; IPR001440; TPR.
DR   Pfam; PF00515; TPR; 2.
SQ   SEQUENCE   623 AA;  68505 MW;  061586EF186F74AF CRC64;

  Query Match              9.1%;  Score 90;  DB 11;  Length 623;
  Best Local Similarity   25.5%;  Pred. No. 4.4;
```

```
  Matches   50;  Conservative   25;  Mismatches   73;  Indels   48;  Gaps    7;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDT 71
              |||:|  ||  :| |     ::|:  :|   |::|                |::|
Db         58 LLLTLQGLPAAVPALPLELTVLCNCIILRASLVQA--------------FTEDLTQDLQ 102

Qy         72 KVNFYAWKRNXSXQQAVE--VWQGLALLSEAVLRGQAL---LVNSSQPWEPLQLHVDKAV 126
              :      :|    |  :|    |||  |  :||   :|   |: :      ||     :
Db        103 R----GLERVLEAQHHLEPKSQQGLKELWHSVLSASSLPPELLPALHCLASLQAVFWMST 158

Qy        127 SGLRSLTTLLRAL-GAQKEAI----------SPPDAASAAPLRTITADTFRKLFRVYSNF 175
                |  || ||: | |:| ::           |||   | |||    |:: | :
Db        159 DHLEDLTLLLQTLNGSQTQSSEDLLLLLKSWSPPAEESPAPLILQDAESLRDV------- 211

Qy        176 LRGKLKLYTGEACRTG 191
                  | |  ||| |
Db        212 ------LLTAFACRQG 221


RESULT 12
Q8Z2M5
ID   Q8Z2M5      PRELIMINARY;      PRT;   346 AA.
AC   Q8Z2M5;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Solute binding receptor protein.
GN   STY3952.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
DR   EMBL; AL627280; CAD03169.1; -.
DR   InterPro; IPR001761; PeriplaBP/Lac1.
DR   Pfam; PF00532; Peripla_BP_like; 1.
KW   Receptor; Complete proteome.
SQ   SEQUENCE   346 AA;  38546 MW;  F80FB1688BC83A8F CRC64;

  Query Match             8.8%;  Score 87.5;  DB 16;  Length 346;
  Best Local Similarity   26.7%;  Pred. No. 3.7;
  Matches   44;  Conservative   22;  Mismatches   48;  Indels   51;  Gaps    9;

Qy         37 RVLERYLLEAKEAENITTG--CAEHCSLNE--NITVPDTKVNFYAWKRNXSXQQAVEVWQ 92
              | | : :||     |:  |   |   :: |  |:| | | |:||           :|::
Db        217 RNLLQEMLERHPDANVVAGSAIAAEAAMGEGRNLTTPLTIVSFYL---------THQVYR 267

Qy         93 GLALLSEAVLRGQALLVNSSQ-PWEPLQLHVDKAVSGLRSLTTLLRALGAQ--KEAISPP 149
              ||        ||  |:  | |  |:           |  ::|  :: |  |   | ||||
Db        268 GLK-------RGHILMALSDQMAWQ----------GELAITQSIKVLQGQPVPENISPP 309

Qy        150 -------DAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEA 187
                    :| ||   |:::   || ::          |  || ||
Db        310 VLILTHNNADSARVRRSLSPPGFRPVY----------LYQYTSEA 344


RESULT 13
Q8ZKZ4
ID   Q8ZKZ4      PRELIMINARY;      PRT;   346 AA.
AC   Q8ZKZ4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Periplasmic sensor in multi-comopnent regulatory system with TorS
DE   (Sensory kinase) and TorR (Regulator), regulates tor operon.
GN   TORT OR STM3825.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
DR   EMBL; AE008878; AAL22684.1; -.
DR   InterPro; IPR001761; PeriplaBP/Lac1.
DR   Pfam; PF00532; Peripla_BP_like; 1.
KW   Complete proteome.
SQ   SEQUENCE   346 AA;  38440 MW;  E37CAB58E49FD716 CRC64;

  Query Match             8.8%;  Score 87.5;  DB 16;  Length 346;
  Best Local Similarity   26.7%;  Pred. No. 3.7;
  Matches   44;  Conservative   22;  Mismatches   48;  Indels   51;  Gaps    9;

Qy         37 RVLERYLLEAKEAENITTG--CAEHCSLNE--NITVPDTKVNFYAWKRNXSXQQAVEVWQ 92
              | | : :||     |:  |   |   :: |  |:| | | |:||           :|::
Db        217 RNLLQEMLERHPDANVVAGSAIAAEAAMGEGRNLTTPLTIVSFYL---------THQVYR 267

Qy         93 GLALLSEAVLRGQALLVNSSQ-PWEPLQLHVDKAVSGLRSLTTLLRALGAQ--KEAISPP 149
              ||        ||  |:  | |  |:           |  ::|  :: |  |   | ||||
Db        268 GLK-------RGHILMALSDQMAWQ----------GELAITQSIKVLQGQPVPENISPP 309

Qy        150 -------DAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEA 187
                    :| ||   |:::   || ::          |  || ||
Db        310 VLILTHNNADSARVRRSLSPPGFRPVY----------LYQYTSEA 344


RESULT 14
Q9D7X0
ID   Q9D7X0      PRELIMINARY;      PRT;   454 AA.
AC   Q9D7X0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   2210018A16Rik protein (Hepatocellular carcinoma-associated antigen
DE   MHCA108).
GN   ZFP296 OR 2210018A16RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Stomach;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
```

```
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Dong X., Su Y., Chen W.;
RT   "MHCA108, a mouse hepatocellular carcinoma-associated antigen.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK008748; BAB25873.1; -.
DR   EMBL; AF450095; AAN76749.1; -.
DR   HSSP; P08153; 1ZFD.
DR   MGD; MGI:1926956; Zfp296.
DR   InterPro; IPR007087; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 5.
DR   SMART; SM00355; ZnF_C2H2; 5.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   454 AA;  48934 MW;  9012C24C9AB02000 CRC64;

  Query Match             8.8%;  Score 87.5;  DB 11;  Length 454;
  Best Local Similarity   24.0%;  Pred. No. 5.2;
  Matches   48;  Conservative   24;  Mismatches   81;  Indels   47;  Gaps    8;

Qy         14 LSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTG-CAEHCSLNENITVPDTK 72
              |   | |||  |   | ||:   |:  :|     |  : :| | | :   |   |   | |
Db         58 LQTASRPLGAPSTCA-PRMPLSSKSSDRQPWTDKHPDLLTCGRCGKIFPLGAIIAFMDHK 116

Qy         73 VNFYAWKRNXSXQQAVEVWQGLALL--SEAVLRGQALLVNS--------SQPWEPL---- 118
                              ||  ||  |: :   : |  |         : ||: |
Db        117 K------------------QGCQLLQVSDPISESKELKAPSCLQCGRQYTSPWKLLCHAQ 158

Qy        119 ---------QLHVD---KAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFR 166
                       |:|     : ||  :   : |:       | |   |:|  |: | |  :
Db        159 WDHGLCIYQTQHLDTPEAPLLGLAEVAAAMSAVAVVAPVESKPPPVSSAARRSPTCDVCK 218

Qy        167 KLFRVYSNFLRGKLKLYTGE 186
              |    :|| |:  :: :|||
Db        219 KTLSSFSN-LKVHMRSHTGE 237


RESULT 15
Q9NB04
ID   Q9NB04      PRELIMINARY;      PRT;   871 AA.
AC   Q9NB04;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Discs LOST.
GN   DLT OR BCDNA:LD22238 OR CG12021.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanentzapf G., Smith C.A., Bhat M., McGlade J., Tepass U.;
RT   "Apical, lateral, and basal polarization cues contribute to the
RT   development of the follicular epithelium during Drosophila
RT   oogenesis.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 4 PDZ/DHR DOMAINS.
DR   EMBL; AF274350; AAF81829.1; -.
DR   HSSP; Q12923; 3PDZ.
DR   FlyBase; FBgn0024246; dlt.
DR   InterPro; IPR004172; L27.
DR   InterPro; IPR001478; PDZ.
DR   Pfam; PF00595; PDZ; 4.
DR   SMART; SM00569; L27; 1.
DR   SMART; SM00228; PDZ; 4.
DR   PROSITE; PS50106; PDZ; 4.
SQ   SEQUENCE   871 AA;  92905 MW;  C9A72F6DB4DB5B4D CRC64;

  Query Match             8.8%;  Score 87;  DB 5;  Length 871;
  Best Local Similarity   25.3%;  Pred. No. 13;
  Matches   57;  Conservative   29;  Mismatches   81;  Indels   58;  Gaps   13;

Qy         13 LLSLLSLP---------------LGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCA 57
              |  ||:||                ||: : |     :|:   |    :     :|:: | |
Db        305 LEELLALPETEKLQVELKKDANGLGITIAG----YVCEKEELSGIFV-----KNVSPGSA 355

Qy         58 EHCSLNENITVPDTKVNFYAWK-RNXSXQQAVEV----WQGLALLSEAVLRG-------Q 105
                 ||  | | |  :       :  |  ||||:     | : |  |  |||        |
Db        356 --ADLNGRIRVNDRIIEVDGQSLQGYSNHQAVELLKKSGQVVNLRLERYLRGPKFEQLQQ 413

Qy        106 ALLVN----SSQPWEPLQLHVDKAVSGLRSLTT-----LLRALGAQKEAISPPDAASAAP 156
              |:  |    || |  | :  :   |:   | ||     : | |  ::||:  |:|    |
Db        414 AIAANDKLPSSAPGTPSRAPMPTPVATTSSATTTPSRSITREL--EEEALPAPEAFMTTP 471

Qy        157 -----LRTITADTF---RKLFRVYSNFLRGKLKLYTGEACRTGDR 193
                  : | |  :|   ::|  |  : | |  |:   |     |:
Db        472 PSVTTMTTTTLSSFGAGKQLVAVRDS-LDGSTKIIPTEVVPLADK 515


Search completed: September  8, 2003, 15:02:05
Job time : 36 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 14:35:11 ; Search time 17 Seconds
(without alignments)
533.891 Million cell updates/sec

Title: US-09-813-775C-34
Perfect score: 991
Sequence: 1 MGVHECPAWLWLLLSLLSLP.....NFLRGKLKLYTGEACRTGDR 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	973	98.2	193	1	EPO_HUMAN	P01588 homo sapien
2	881.5	89.0	192	1	EPO_MACFA	P07865 macaca fasc
3	879.5	88.7	192	1	EPO_MACMU	Q28513 macaca mula
4	778.5	78.6	192	1	EPO_FELCA	P33708 felis silve
5	762.5	76.9	192	1	EPO_RAT	P29676 rattus norv
6	752.5	75.9	194	1	EPO_SHEEP	P33709 ovis aries
7	748.5	75.5	190	1	EPO_PIG	P49157 sus scrofa
8	746.5	75.3	192	1	EPO_MOUSE	P07321 mus musculu
9	740.5	74.7	192	1	EPO_BOVIN	P48617 bos taurus
10	699.5	70.6	175	1	EPO_CANFA	P33707 canis famil
11	122.5	12.4	352	1	TPO_CANFA	P42705 canis famil
12	98.5	9.9	353	1	TPO_HUMAN	P40225 homo sapien
13	91	9.2	356	1	TPO_MOUSE	P40226 mus musculu
14	87.5	8.8	326	1	TPO_RAT	P49745 rattus norv
15	86.5	8.7	622	1	FACG_HUMAN	O15287 homo sapien
16	84	8.5	339	1	MURB_PSEAE	Q9hzm7 pseudomonas
17	80	8.1	353	1	NADA_YERPE	Q8zgy8 yersinia pe
18	80	8.1	1980	1	MY9B_RAT	Q63358 rattus norv
19	79	8.0	1089	1	IMB3_YEAST	P32337 saccharomyc
20	76.5	7.7	381	1	APA_MYCAV	Q48919 mycobacteri
21	76.5	7.7	543	1	CH60_BARBA	P35635 bartonella
22	76	7.7	552	1	CH60_COXBU	P19421 coxiella bu
23	75	7.6	263	1	YH25_DEIRA	O83030 deinococcus
24	75	7.6	778	1	RGL2_MOUSE	Q61193 mus musculu
25	74.5	7.5	475	1	Z342_HUMAN	Q8wuu4 homo sapien
26	74.5	7.5	3033	1	POLG_HCVJ8	P26661 h genome po
27	73.5	7.4	388	1	TRA6_BURCE	P24575 burkholderi
28	73.5	7.4	897	1	EP15_MOUSE	P42567 mus musculu
29	72.5	7.3	347	1	NADA_SALTI	Q8z8b8 salmonella
30	71.5	7.2	220	1	Y085_MYCTU	Q10882 mycobacteri
31	71.5	7.2	283	1	AROE_XANAC	Q8pfe9 xanthomonas
32	71.5	7.2	486	1	BAF1_KLULA	P26375 kluyveromyc
33	71.5	7.2	907	1	GACS_PSESY	P48027 pseudomonas
34	71.5	7.2	3164	1	TEGU_HSV11	P10220 herpes simp
35	71	7.2	224	1	MER1_HUMAN	Q9um22 homo sapien
36	71	7.2	342	1	TORT_ECO57	P58358 escherichia
37	71	7.2	342	1	TORT_ECOLI	P38683 escherichia
38	71	7.2	402	1	YYCB_BACSU	P37482 bacillus su
39	71	7.2	547	1	G6P2_NEIMA	Q9jss6 neisseria m
40	71	7.2	926	1	ATCL_SYNP7	P37278 synechococc
41	70.5	7.1	279	1	LEP4_ERWCA	P31712 erwinia car
42	70.5	7.1	283	1	AROE_XANCP	Q8p3w6 xanthomonas
43	70.5	7.1	1402	1	N160_MOUSE	Q9z0w3 mus musculu
44	70	7.1	347	1	NADA_SALTY	P24519 salmonella
45	70	7.1	473	1	RT4R_HUMAN	Q9bzr6 homo sapien

ALIGNMENTS

RESULT 1
EPO_HUMAN
ID EPO_HUMAN STANDARD; PRT; 193 AA.
AC P01588; Q9UDZ0; Q9UEZ5; Q9UHA0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Erythropoietin precursor (Epoetin).
GN EPO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85137899; PubMed=3838366;
RA Jacobs K., Shoemaker C., Rudersdorf R., Neill S.D., Kaufman R.J.,
RA Mufson A., Seehra J., Jones S.S., Hewick R., Fritsch E.F.,
RA Kawakita M., Shimizu T., Miyake T.;
RT "Isolation and characterization of genomic and cDNA clones of human
RT erythropoietin.";
RL Nature 313:806-810(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067948; PubMed=3865178;
RA Lin F.-K., Suggs S., Lin C.-H., Browne J.K., Smalling R., Egrie J.C.,
RA Chen K.K., Fox G.M., Martin F., Stabinsky Z., Badrawi S.M., Lai P.-H.,
RA Goldwasser E.;
RT "Cloning and expression of the human erythropoietin gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7580-7584(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99018118; PubMed=9799793;
RA Gloeckner G., Scherer S., Schattevoy R., Boright A.P., Weber J.,
RA Tsui L.-C., Rosenthal A.;
RT "Large-scale sequencing of two regions in human chromosome 7q22:
RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT reveals 17 genes.";
RL Genome Res. 8:1060-1073(1998).
RN [4]
RP SEQUENCE FROM N.A.
RA Rupert J.L., Hochachka P.W.;
RT "Erythropoietin gene sequence in the Quechua, a high altitude native
RT population.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE OF 58-193 FROM N.A., AND VARIANTS HEPATOCELLULAR CARCINOMA
RP 131-ASN-PHE-132 AND GLN-149.
RX MEDLINE=93384593; PubMed=8396923;
RA Funakoshi A., Muta H., Baba T., Shimizu S.;
RT "Gene expression of mutant erythropoietin in hepatocellular
RT carcinoma.";
RL Biochem. Biophys. Res. Commun. 195:717-722(1993).
RN [6]
RP SEQUENCE OF 28-193, AND DISULFIDE BONDS.
RC TISSUE=Urine;

```
RX   MEDLINE=86140080; PubMed=3949763;
RA   Lai P.H., Everett R., Wang F.F., Arakawa T., Goldwasser E.;
RT   "Structural characterization of human erythropoietin.";
RL   J. Biol. Chem. 261:3116-3121(1986).
RN   [7]
RP   PRELIMINARY SEQUENCE OF 28-57.
RX   MEDLINE=84135751; PubMed=6698989;
RA   Yanagawa S., Hirade K., Ohnota H., Sasaki R., Chiba H., Ueda M.,
RA   Goto M.;
RT   "Isolation of human erythropoietin with monoclonal antibodies.";
RL   J. Biol. Chem. 259:2707-2710(1984).
RN   [8]
RP   STRUCTURE OF CARBOHYDRATES.
RX   MEDLINE=88153657; PubMed=3346214;
RA   Takeuchi M., Takasaki S., Miyazaki H., Kato T., Hoshi S., Kochibe N.,
RA   Kobata A.;
RT   "Comparative study of the asparagine-linked sugar chains of human
RT   erythropoietins purified from urine and the culture medium of
RT   recombinant Chinese hamster ovary cells.";
RL   J. Biol. Chem. 263:3657-3663(1988).
RN   [9]
RP   STRUCTURE OF CARBOHYDRATES.
RX   MEDLINE=89118279; PubMed=3219367;
RA   Sasaki H., Ochi N., Dell A., Fukuda M.;
RT   "Site-specific glycosylation of human recombinant erythropoietin:
RT   analysis of glycopeptides or peptides at each glycosylation site by
RT   fast atom bombardment mass spectrometry.";
RL   Biochemistry 27:8618-8626(1988).
RN   [10]
RP   STRUCTURE OF CARBOHYDRATES.
RX   MEDLINE=92314463; PubMed=1820196;
RA   Takeuchi M., Kobata A.;
RT   "Structures and functional roles of the sugar chains of human
RT   erythropoietins.";
RL   Glycobiology 1:337-346(1991).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=98445092; PubMed=9774108;
RA   Syed R.S., Reid S.W., Li C., Cheetham J.C., Aoki K.H., Liu B.,
RA   Zhan H., Osslund T.D., Chirino A.J., Zhang J., Finer-Moore J.,
RA   Elliott S., Sitney K., Katz B.A., Matthews D.J., Wendoloski J.J.,
RA   Egrie J., Stroud R.M.;
RT   "Efficiency of signalling through cytokine receptors depends
RT   critically on receptor orientation.";
RL   Nature 395:511-516(1998).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- PHARMACEUTICAL: Used for the treatment of anemia. Available under
CC       the names Epogen (Amgen), Epogin (Chugai), Epomax (Elanex), Eprex
CC       (Janssen-Cilag), NeoRecormon or Recormon (Roche), and Procrit
CC       (Ortho Biotech). Variations in the glycosylation pattern of EPO
CC       distinguishes these products. Epogen, Epogin, Eprex and Procrit
CC       are generically known as epoetin alfa, NeoRecormon and Recormon as
CC       epoetin beta and Epomax as epoetin omega.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   -!- DATABASE: NAME=R&D Systems' cytokine source book: EPO;
CC       WWW="http://www.rndsystems.com/asp/g_sitebuilder.asp?bodyId=197".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X02158; CAA26095.1; -.
DR   EMBL; X02157; CAA26094.1; -.
DR   EMBL; M11319; AAA52400.1; -.
DR   EMBL; AF053356; AAC78791.1; -.
DR   EMBL; AF202308; AAF23132.1; -.
DR   EMBL; AF202306; AAF23132.1; JOINED.
DR   EMBL; AF202307; AAF23132.1; JOINED.
DR   EMBL; AF202310; AAF23133.1; -.
DR   EMBL; AF202309; AAF23133.1; JOINED.
DR   EMBL; AF202311; AAF17572.1; -.
DR   EMBL; AF202314; AAF23134.1; -.
DR   EMBL; AF202312; AAF23134.1; JOINED.
DR   EMBL; AF202313; AAF23134.1; JOINED.
DR   EMBL; S65458; AAD13964.1; -.
DR   PIR; A01855; ZUHU.
DR   PDB; 1EER; 01-OCT-99.
DR   PDB; 1CN4; 11-AUG-99.
DR   PDB; 1BUY; 10-SEP-99.
DR   GlycoSuiteDB; P01588; -.
DR   Genew; HGNC:3415; EPO.
DR   MIM; 133170; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0006950; P:response to stress; TAS.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal; Pharmaceutical;
KW   3D-structure; Polymorphism.
FT   SIGNAL       1    27
FT   CHAIN       28   193       ERYTHROPOIETIN.
FT   PROPEP     190   193       MAY BE REMOVED IN PROCESSED PROTEIN.
FT   DISULFID    34   188
FT   DISULFID    56    60
FT   CARBOHYD    51    51       N-LINKED (GLCNAC. . .).
FT                               /FTId=CAR_000052.
FT   CARBOHYD    65    65       N-LINKED (GLCNAC. . .).
FT                               /FTId=CAR_000166.
FT   CARBOHYD   110   110       N-LINKED (GLCNAC. . .).
FT                               /FTId=CAR_000192.
FT   CARBOHYD   153   153       O-LINKED (GALNAC. . .).
FT   VARIANT    131   132       SL -> NF (IN AN HEPATOCELLULAR
FT                               CARCINOMA).
FT                               /FTId=VAR_009870.
FT   VARIANT    149   149       P -> Q (IN AN HEPATOCELLULAR CARCINOMA).
FT                               /FTId=VAR_009871.
FT   CONFLICT    40    40       E -> Q (IN REF. 1; CAA26095).
FT   CONFLICT    85    85       Q -> QQ (IN REF. 5).
FT   CONFLICT   140   140       G -> R (IN REF. 1; CAA26095).
FT   HELIX       32    34
FT   HELIX       36    52
FT   HELIX       53    55
FT   TURN        57    58
FT   STRAND      61    68
FT   STRAND      73    73
FT   HELIX       75    78
FT   TURN        79    80
FT   HELIX       83   109
FT   HELIX      118   138
FT   TURN       139   140
FT   HELIX      141   147
FT   TURN       148   149
FT   STRAND     160   164
FT   HELIX      165   177
FT   TURN       178   178
FT   HELIX      179   188
SQ   SEQUENCE   193 AA;  21306 MW;  C91F0E4C26A52033 CRC64;

  Query Match             98.2%;  Score 973;  DB 1;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 3.5e-85;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
```

```
Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193
```

RESULT 2
EPO_MACFA
ID EPO_MACFA STANDARD; PRT; 192 AA.
AC P07865;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87055236; PubMed=2877922;
RA Lin F.-K., Lin C.-H., Lai P.-H., Browne J.K., Egrie J.C., Smalling R.,
RA Fox G.M., Chen K.K., Castro M., Suggs S.;
RT "Monkey erythropoietin gene: cloning, expression and comparison with
RT the human erythropoietin gene.";
RL Gene 44:201-209(1986).
CC -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; M18189; AAA36841.1; -.
DR PIR; JQ0173; JQ0173.
DR HSSP; P01588; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythroptn.
DR Pfam; PF00758; EPO_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO; 1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 192 ERYTHROPOIETIN.
FT DISULFID 34 187 BY SIMILARITY.
FT DISULFID 56 60 BY SIMILARITY.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 152 152 O-LINKED (GALNAC. . .) (BY SIMILARITY).
SQ SEQUENCE 192 AA; 21113 MW; E8A900F442AD4522 CRC64;

Query Match 89.0%; Score 881.5; DB 1; Length 192;
Best Local Similarity 89.6%; Pred. No. 1.7e-76;

Matches 173; Conservative 8; Mismatches 11; Indels 1; Gaps 1;

```
Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              |||||||||||||||:|||||||| |||||||||||||||||||||:| ||:| |
Db          1 MGVHECPAWLWLLLSLVSLPLGLPVPGAPPRLICDSRVLERYLLEAKEAENVTMGCSESC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||    ||||||||||||||||||||:| |||||:|||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQAVLANSSQPFEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              |:|||:|||||:|||||||||| |||| ||||||||||||||| ||||||||||||
Db        121 HMDKAISGLRSITTLLRALGAQ-EAISLPDAASAAPLRTITADTFCKLFRVYSNFLRGKL 179

Qy        181 KLYTGEACRTGDR 193
              ||||||||| |||
Db        180 KLYTGEACRRGDR 192
```

RESULT 3
EPO_MACMU
ID EPO_MACMU STANDARD; PRT; 192 AA.
AC Q28513;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=93372347; PubMed=8364201;
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993).
CC -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; L10609; AAA36842.1; -.
DR PIR; I84613; I84613.
DR HSSP; P01588; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythroptn.
DR Pfam; PF00758; EPO_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO; 1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 192 ERYTHROPOIETIN.
FT DISULFID 34 187 BY SIMILARITY.
FT DISULFID 56 60 BY SIMILARITY.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (BY SIMILARITY).

```
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD    152    152       O-LINKED (GALNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   192 AA;  21081 MW;  275560A264628CD1 CRC64;

  Query Match             88.7%;  Score 879.5;  DB 1;  Length 192;
  Best Local Similarity   89.1%;  Pred. No. 2.6e-76;
  Matches  172;  Conservative    9;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              |||||||||||||||:|||||||| |||||:|||||||||||||||||:| ||:| |
Db          1 MGVHECPAWLWLLLSLVSLPLGLPVPGAPPRLVCDSRVLERYLLEAKEAENVTMGCSESC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||    ||||||||||||||||||||:| |||||:|||||
Db         61 SLNENITVPDTKVNFYAWKRIEVGQQAVEVWQGLALLSEAVLRGQAVLANSSQPFEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              |:|||:|||||:|||||||||| |||| |||||||||||||||| ||||||||||||||
Db        121 HMDKAISGLRSITTLLRALGAQ-EAISLPDAASAAPLRTITADTFCKLFRVYSNFLRGKL 179

Qy        181 KLYTGEACRTGDR 193
              ||||||||| |||
Db        180 KLYTGEACRRGDR 192


RESULT 4
EPO_FELCA
ID   EPO_FELCA      STANDARD;      PRT;   192 AA.
AC   P33708;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Goodman R.E., Bell R.G.;
RL   Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 5-192 FROM N.A.
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA   Czelusniak J., Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U00685; AAA18282.1; -.
DR   EMBL; L10606; AAA30807.1; -.
DR   PIR; I46083; I46083.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    192       ERYTHROPOIETIN.
FT   DISULFID    33    187       BY SIMILARITY.
FT   DISULFID    55     59       BY SIMILARITY.
FT   CARBOHYD    50     50       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   109    109       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    44     44       G -> E (IN REF. 2).
SQ   SEQUENCE   192 AA;  20914 MW;  61C5EA0F5E937293 CRC64;

  Query Match             78.6%;  Score 778.5;  DB 1;  Length 192;
  Best Local Similarity   81.9%;  Pred. No. 9.8e-67;
  Matches  158;  Conservative    8;  Mismatches   26;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||  |||| | |||||| |||||||||||||||||||||||:| |:|||||:| |||| |
Db          1 MGSCECPA-LLLLLSLLLLPLGLPVLGAPPRLICDSRVLERYILGAREAENVTMGCAEGC 59

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              | :||||||||||||| |||    |||||||||||||||:||||||| ||||| | |||
Db         60 SFSENITVPDTKVNFYTWKRMDVGQQAVEVWQGLALLSEAILRGQALLANSSQPSETLQL 119

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||| |||||:|||||||||| | |:| |||||| | ||  ||||:|||||||||
Db        120 HVDKAVSSLRSLTSLLRALGAQKEATSLPEATSAAPLRTFTVDTLCKLFRIYSNFLRGKL 179

Qy        181 KLYTGEACRTGDR 193
              |||||||| |||
Db        180 TLYTGEACRRGDR 192


RESULT 5
EPO_RAT
ID   EPO_RAT        STANDARD;      PRT;   192 AA.
AC   P29676; P70504;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Kidney;
RX   MEDLINE=93042015; PubMed=1420369;
RA   Nagao M., Suga H., Okano M., Masuda S., Narita H., Ikura K.,
RA   Sasaki R.;
RT   "Nucleotide sequence of rat erythropoietin.";
RL   Biochim. Biophys. Acta 1171:99-102(1992).
RN   [2]
RP   SEQUENCE OF 4-192 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Mulcahy L.S., Czelusniak J.,
RA   Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D10763; BAA01593.1; -.
DR   EMBL; L10608; AAA41126.1; -.
DR   PIR; S28148; S28148.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    192       ERYTHROPOIETIN.
FT   DISULFID    33    187       BY SIMILARITY.
FT   CARBOHYD    50     50       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD    64     64       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD   109    109       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   192 AA;  21286 MW;  3EA632737E7D2443 CRC64;

  Query Match             76.9%;  Score 762.5;  DB 1;  Length 192;
  Best Local Similarity   79.8%;  Pred. No. 3.2e-65;
  Matches  154;  Conservative   13;  Mismatches   25;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||| | |  | |||||| :||||||| ||||||||||||||||:|||||||:| ||||
Db          1 MGVPERPT-LLLLLSLLLIPLGLPVLCAPPRLICDSRVLERYILEAKEAENVTMGCAEGP 59

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
               |:|||||||||||||||||    :|||||||||:|||||:|: |||  ||||| | |||
Db         60 RLSENITVPDTKVNFYAWKRMKVEEQAVEVWQGLSLLSEAILQAQALQANSSQPPESLQL 119

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              |:|||:|||||||:||| |||||| :|||||  |||||||:||||| ||||||||||||||
Db        120 HIDKAISGLRSLTSLLRVLGAQKELMSPPDATQAAPLRTLTADTFCKLFRVYSNFLRGKL 179

Qy        181 KLYTGEACRTGDR 193
              ||||||||| |||
Db        180 KLYTGEACRRGDR 192


RESULT 6
EPO_SHEEP
ID   EPO_SHEEP      STANDARD;      PRT;   194 AA.
AC   P33709; Q28572;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=93351736; PubMed=8349021;
RA   Fu P., Evans B., Lim G.B., Moritz K., Wintour M.E.;
RT   "The sheep erythropoietin gene: molecular cloning and effect of
RT   hemorrhage on plasma erythropoietin and renal/liver messenger RNA in
RT   adult sheep.";
RL   Mol. Cell. Endocrinol. 93:107-116(1993).
RN   [2]
RP   SEQUENCE OF 4-194 FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA   Czelusniak J., Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z24681; CAA80848.1; -.
DR   EMBL; L10610; AAA31518.1; -.
DR   PIR; I46401; I46401.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     27       BY SIMILARITY.
FT   CHAIN       28    194       ERYTHROPOIETIN.
FT   DISULFID    34    189       BY SIMILARITY.
FT   DISULFID    56     60       BY SIMILARITY.
FT   CARBOHYD    51     51       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    65     65       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    16     16       F -> L (IN REF. 2).
FT   CONFLICT   108    108       L -> P (IN REF. 2).
SQ   SEQUENCE   194 AA;  21335 MW;  C025AAB0528131A9 CRC64;

  Query Match             75.9%;  Score 752.5;  DB 1;  Length 194;
  Best Local Similarity   77.8%;  Pred. No. 2.9e-64;
  Matches  151;  Conservative   10;  Mismatches   32;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||  :|   | |||| |  |||||||||||||||||||||:|||:|||| | |||| |
Db          1 MGARDCTPLLLLLLSFLLFPLGLPVLGAPPRLICDSRVLERYILEAREAENATMGCAEGC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
               | :|||||||||||||||||    |||:||||||||||||: ||||||  |:||| | |:|
Db         61 SFSENITVPDTKVNFYAWKRMEVQQQALEVWQGLALLSEAIFRGQALLANASQPCEALRL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAA-SAAPLRTITADTFRKLFRVYSNFLRGK 179
              ||||||||||||:||||||||||||  |||  ||||||  | |   ||||:|||||||
Db        121 HVDKAVSGLRSLTSLLRALGAQKEAIPLPDATPSAAPLRIFTVDALSKLFRIYSNFLRGK 180

Qy        180 LKLYTGEACRTGDR 193
              | |||||||| |||
Db        181 LTLYTGEACRRGDR 194


RESULT 7
EPO_PIG
ID   EPO_PIG        STANDARD;      PRT;   190 AA.
AC   P49157;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor (Fragment).
GN   EPO.
```

```
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA   Czelusniak J., Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10607; AAA31029.1; -.
DR   PIR; I46578; I46578.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     22       POTENTIAL.
FT   CHAIN        23    190       ERYTHROPOIETIN.
FT   DISULFID     29    185       BY SIMILARITY.
FT   DISULFID     51     55       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    105    105       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   190 AA;  20888 MW;  A75BD6CCE5077E2A CRC64;

  Query Match             75.5%;  Score 748.5;  DB 1;  Length 190;
  Best Local Similarity   81.2%;  Pred. No. 6.8e-64;
  Matches  155;  Conservative    7;  Mismatches   26;  Indels    3;  Gaps    2;

Qy          5 ECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNE 64
              |||| | |||||| |||||||||||||||||||||||||:||||| || | |||| || :|
Db          1 ECPARL-LLLSLLLLPLGLPVLGAPPRLICDSRVLERYILEAKEGENATMGCAESCSFSE 59

Qy         65 NITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDK 124
              |||||||||||||||    |||:|||||||||||:|:||||| ||||| | |||||||
Db         60 NITVPDTKVNFYAWKRMEVQQQAMEVWQGLALLSEAILQGQALLANSSQPSEALQLHVDK 119

Qy        125 AVSGLRSLTTLLRALGAQKEAISPPDA--ASAAPLRTITADTFRKLFRVYSNFLRGKLKL 182
              |||||||||:|||||||||||||  |||  :|| ||||   ||  |||| |||||||| |
Db        120 AVSGLRSLTSLLRALGAQKEAIPLPDASPSSATPLRTFAVDTLCKLFRNYSNFLRGKLTL 179

Qy        183 YTGEACRTGDR 193
              |||||||  ||
Db        180 YTGEACRRRDR 190


RESULT 8
EPO_MOUSE
ID   EPO_MOUSE      STANDARD;      PRT;   192 AA.
AC   P07321;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87039105; PubMed=3773894;
RA   Shoemaker C.B., Mitsock L.D.;
RT   "Murine erythropoietin gene: cloning, expression, and human gene
RT   homology.";
RL   Mol. Cell. Biol. 6:849-858(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87039104; PubMed=3022133;
RA   McDonald J.D., Lin F.-K., Goldwasser E.;
RT   "Cloning, sequencing, and evolutionary analysis of the mouse
RT   erythropoietin gene.";
RL   Mol. Cell. Biol. 6:842-848(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RX   MEDLINE=21138439; PubMed=11239002;
RA   Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,
RA   Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
RA   Miller W., Koop B.F.;
RT   "Comparative analysis of the gene-dense ACHE/TFR2 region on human
RT   chromosome 7q22 with the orthologous region on mouse chromosome 5.";
RL   Nucleic Acids Res. 29:1352-1365(2001).
RN   [4]
RP   SEQUENCE OF 1-52 FROM N.A.
RC   STRAIN=ICFW;
RX   MEDLINE=98030528; PubMed=9365246;
RA   Chretien S., Duprez V., Maouche L., Gisselbrecht S., Mayeux P.,
RA   Lacombe C.;
RT   "Abnormal erythropoietin (Epo) gene expression in the murine
RT   erythroleukemia IW32 cells results from a rearrangement between the
RT   G- protein beta2 subunit gene and the Epo gene.";
RL   Oncogene 15:1995-1999(1997).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M12482; AAA37568.1; -.
DR   EMBL; M12930; AAA37570.1; -.
DR   EMBL; AF312033; AAK28825.1; -.
DR   EMBL; Y11971; CAA72707.1; -.
DR   PIR; A24902; A24902.
DR   HSSP; P01588; 1CN4.
DR   MGD; MGI:95407; Epo.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
```

```
FT   SIGNAL        1     26
FT   CHAIN        27    192       ERYTHROPOIETIN.
FT   DISULFID     33    187       BY SIMILARITY.
FT   CARBOHYD     50     50       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD     64     64       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD    109    109       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   192 AA;  21365 MW;  65F94E214E0DEF2E CRC64;

  Query Match             75.3%;  Score 746.5;  DB 1;  Length 192;
  Best Local Similarity   77.2%;  Pred. No. 1.1e-63;
  Matches  149;  Conservative   15;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||| | |  | |||||| :||||||| ||||||||||||||:||||||||:| ||||
Db          1 MGVPERPT-LLLLLSLLLIPLGLPVLCAPPRLICDSRVLERYILEAKEAENVTMGCAEGP 59

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
               |:|||||||||||||||||    :||:||||||:|||||:|: |||| ||||| | |||
Db         60 RLSENITVPDTKVNFYAWKRMEVEEQAIEVWQGLSLLSEAILQAQALLANSSQPPETLQL 119

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              |:|||:|||||||:||| |||||| :|||||    |||||:| ||| ||||||:|||||||
Db        120 HIDKAISGLRSLTSLLRVLGAQKELMSPPDTTPPAPLRTLTVDTFCKLFRVYANFLRGKL 179

Qy        181 KLYTGEACRTGDR 193
              |||||| || |||
Db        180 KLYTGEVCRRGDR 192


RESULT 9
EPO_BOVIN
ID   EPO_BOVIN      STANDARD;      PRT;   192 AA.
AC   P48617;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Boran; TISSUE=Kidney;
RX   MEDLINE=96257233; PubMed=8666286;
RA   Suliman H.B., Majiwa P.A.O., Feldman B.F., Mertens B.,
RA   Logan-Henfrey L.L.;
RT   "Cloning of a cDNA encoding bovine erythropoietin and analysis of its
RT   transcription in selected tissues.";
RL   Gene 171:275-280(1996).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L41354; AAB41268.1; -.
DR   EMBL; U44762; AAA86653.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    192       ERYTHROPOIETIN.
FT   DISULFID     32    187       BY SIMILARITY.
FT   DISULFID     54     58       BY SIMILARITY.
FT   CARBOHYD     49     49       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    108    108       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   192 AA;  21075 MW;  DBC419022F7B483A CRC64;

  Query Match             74.7%;  Score 740.5;  DB 1;  Length 192;
  Best Local Similarity   77.3%;  Pred. No. 4e-63;
  Matches  150;  Conservative   10;  Mismatches   31;  Indels    3;  Gaps    2;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||  :|     |:|| |  ||| |||||| |||||||||||:|||:|||| | |||| |
Db          1 MGARDCTP--LLMLSFLLFPLGFPVLGAPARLICDSRVLERYILEAREAENATMGCAEGC 58

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              | |||||||||||||||||    |||:|||||||||||:|||||| |:||| | |:|
Db         59 SFNENITVPDTKVNFYAWKRMEVQQQALEVWQGLALLSEAILRGQALLANASQPCEALRL 118

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAA-SAAPLRTITADTFRKLFRVYSNFLRGK 179
              |||||||||||||:|||||||||||||| |||  ||||||  | |   ||||:||||||||
Db        119 HVDKAVSGLRSLTSLLRALGAQKEAISLPDATPSAAPLRAFTVDALSKLFRIYSNFLRGK 178

Qy        180 LKLYTGEACRTGDR 193
              | |||||||| |||
Db        179 LTLYTGEACRRGDR 192


RESULT 10
EPO_CANFA
ID   EPO_CANFA      STANDARD;      PRT;   175 AA.
AC   P33707;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor (Fragment).
GN   EPO.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA   Czelusniak J., Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L13027; AAA30842.1; -.
DR   PIR; I46199; I46199.
```

```
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     22       BY SIMILARITY.
FT   CHAIN        23   >175       ERYTHROPOIETIN.
FT   DISULFID     29   >175       BY SIMILARITY.
FT   DISULFID     51     55       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    105    105       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     175    175
SQ   SEQUENCE   175 AA;  19193 MW;  B504F8DE86676BF4 CRC64;

  Query Match             70.6%;  Score 699.5;  DB 1;  Length 175;
  Best Local Similarity   80.1%;  Pred. No. 2.7e-59;
  Matches  141;  Conservative   12;  Mismatches   22;  Indels    1;  Gaps    1;

Qy          5 ECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNE 64
              |||| | |||||| |||||||||||||||||||||||:|||:||||:| |||: || :|
Db          1 ECPA-LLLLLSLLLPLGLPVLGAPPRLICDSRVLERYILEAREAENVTMGCAQGCSFSE 59

Qy         65 NITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDK 124
              |||||||||||| |||    |||:||||||||||||:||||| |:||| |  ||||||
Db         60 NITVPDTKVNFYTWKRMDVGQQALEVWQGLALLSEAILRGQALLANASQPSETPQLHVDK 119

Qy        125 AVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||| |||||:|||||||||||:| |: || ||||| | ||  ||||:|||||||||
Db        120 AVSSLRSLTSLLRALGAQKEAMSLPEEASPAPLRTFTVDTLCKLFRIYSNFLRGKL 175


RESULT 11
TPO_CANFA
ID   TPO_CANFA      STANDARD;      PRT;   352 AA.
AC   P42705;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE   (C-MPL ligand) (ML) (Megakaryocyte growth and development factor)
DE   (MGDF).
GN   THPO OR TPO.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 24-44.
RC   TISSUE=Kidney;
RX   MEDLINE=94291201; PubMed=8020099;
RA   Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
RA   Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
RA   Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
RA   Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangelinan M.,
RA   Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
RA   Shutter J., Chute H., Basu R., Selander L., Trollinger D., Sieu L.,
RA   Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA   Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
RA   Pacifici R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
RA   Bosselman R.A.;
RT   "Identification and cloning of a megakaryocyte growth and development
RT   factor that is a ligand for the cytokine receptor Mpl.";
RL   Cell 77:1117-1117(1994).
CC   -!- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
CC       AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
CC       CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
CC       MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC       TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003978; thrombopoeitin.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR01485; THROMBOPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Cytokine; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    352       THROMBOPOIETIN.
FT   DISULFID    28    172       POTENTIAL.
FT   DISULFID    50    106       POTENTIAL.
FT   CARBOHYD   185    185       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   206    206       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   255    255       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   332    332       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   347    347       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   352 AA;  37641 MW;  024F3B41B061FBD8 CRC64;

  Query Match             12.4%;  Score 122.5;  DB 1;  Length 352;
  Best Local Similarity   25.5%;  Pred. No. 0.00034;
  Matches   47;  Conservative   24;  Mismatches   70;  Indels   43;  Gaps    6;

Qy         12 LLLSLLSLPLGL-PVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPD 70
              ||: :| |   | | | |||   || |:| : | ::    :  : | :   |:  : :|
Db          7 LLVVMLLLTARLDPCLPAPP--ACDPRLLNKMLRDSHVLHSRLSQCPDIYPLSTPVLLPA 64

Qy         71 TKVNFYAWKRNXSXQQAVEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSG 128
                 :   ||      :| :||  :||| : ||  |||                      |
Db         65 VDFSLGEWKTQKEQTKAQDVWGAVALLLDGVLAARGQL--------------------G 103

Qy        129 LRSLTTLLRALGAQKEAI-----------SPPDAASAAPLRTITADTFRKLFRVYSNFLR 177
                |::||  |  |   :            ||        || |      :|  :   ||
Db        104 PSCLSSLLGQLSGQVRLLLGALQGLLGTQLPPQG------RTTTHKDPNAIFLSFQQLLR 157

Qy        178 GKLK 181
              ||::
Db        158 GKVR 161


RESULT 12
TPO_HUMAN
ID   TPO_HUMAN      STANDARD;      PRT;   353 AA.
AC   P40225; Q13020; Q15790; Q15791; Q15792;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE   (Myeloproliferative leukemia virus oncogene ligand) (C-mpl ligand)
DE   (ML) (Megakaryocyte growth and development factor) (MGDF).
GN   THPO.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal liver;
RX   MEDLINE=94261202; PubMed=8202154;
RA   de Sauvage F.J., Hass P.E., Spencer S.D., Malloy B.E., Gurney A.L.,
RA   Spencer S.A., Darbonne W.C., Henzel W.J., Wong S.C., Kuang W.-J.,
RA   Oles K.J., Hultgren B., Solberg L.A. Jr., Goeddel D.V., Eaton D.L.;
RT   "Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl
RT   ligand.";
RL   Nature 369:533-538(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal liver;
RX   MEDLINE=94291201; PubMed=8020099;
RA   Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
RA   Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
```

```
RA   Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
RA   Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangelinan M.,
RA   Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
RA   Shutter J., Chute H., Basu R., Selander L., Trollinger D., Sieu L.,
RA   Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA   Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
RA   Pacifici R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
RA   Bosselman R.A.;
RT   "Identification and cloning of a megakaryocyte growth and development
RT   factor that is a ligand for the cytokine receptor Mpl.";
RL   Cell 77:1117-1124(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=95108091; PubMed=7809166;
RA   Foster D.C., Sprecher C.A., Grant F.J., Kramer J.M., Kuijper J.L.,
RA   Holly R.D., Whitmore T.E., Heipel M.D., Bell L.A.N., Ching A.F.,
RA   McGrane V., Hart C., O'Hara P.J., Lok S.;
RT   "Human thrombopoietin: gene structure, cDNA sequence, expression, and
RT   chromosomal localization.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:13023-13027(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=95010765; PubMed=7926023;
RA   Sohma Y., Akahori H., Seki N., Hori T.-A., Ogami K., Kawamura K.,
RA   Miyazaki H.;
RT   "Molecular cloning and chromosomal localization of the human
RT   thrombopoietin gene.";
RL   FEBS Lett. 353:57-61(1994).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX   MEDLINE=95152076; PubMed=7849319;
RA   Gurney A.L., Kuang W.-J., Xie M.-H., Malloy B.E., Eaton D.L.,
RA   de Sauvage F.J.;
RT   "Genomic structure, chromosomal localization, and conserved
RT   alternative splice forms of thrombopoietin.";
RL   Blood 85:981-988(1995).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Liver;
RX   MEDLINE=96015174; PubMed=8537317;
RA   Kato T., Ogami K., Shimada Y., Iwamatsu A., Sohma Y., Akahori H.,
RA   Horie K., Kokubo A., Kudo Y., Maeda E., Kobayashi K., Ohashi H.,
RA   Ozawa T., Inoue H., Kawamura K., Miyazaki H.;
RT   "Purification and characterization of thrombopoietin.";
RL   J. Biochem. 118:229-236(1995).
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RX   MEDLINE=95122483; PubMed=7822271;
RA   Chang M., McNinch J., Basu R., Shutter J., Hsu R., Perkins C., Mar V.,
RA   Suggs S., Welcher A., Li L., Lu H., Bartley T., Hunt P., Martin F.,
RA   Samal B., Bogenberger J.;
RT   "Cloning and characterization of the human megakaryocyte growth and
RT   development factor (MGDF) gene.";
RL   J. Biol. Chem. 270:511-514(1995).
RN   [8]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RA   Im S.H., Lee W.S., Chung K.H.;
RT   "Cloning and sequencing of human thrombopoietin.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
CC       AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
CC       CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
CC       MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=1;
CC         IsoId=P40225-1; Sequence=Displayed;
CC       Name=2; Synonyms=TPO-2;
CC         IsoId=P40225-2; Sequence=VSP_001450;
CC       Name=3; Synonyms=Truncated;
CC         IsoId=P40225-3; Sequence=VSP_001451;
CC   -!- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC       TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   -!- DATABASE: NAME=R&D Systems' cytokine source book: TPO;
CC       WWW="http://www.rndsystems.com/asp/g_sitebuilder.asp?bodyId=225".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L33410; AAA59857.1; -.
DR   EMBL; U11025; AAA50553.1; -.
DR   EMBL; L36051; AAC37568.1; -.
DR   EMBL; L36052; AAC37566.1; -.
DR   EMBL; D32046; BAA06807.1; -.
DR   EMBL; S76771; AAB33390.1; -.
DR   EMBL; D32047; BAA21930.1; -.
DR   EMBL; U59493; AAB03392.1; -.
DR   EMBL; U59494; AAB03393.1; -.
DR   EMBL; U59495; AAB03394.1; -.
DR   EMBL; U17071; AAA74083.1; -.
DR   PIR; I59281; I80105.
DR   Genew; HGNC:11795; THPO.
DR   MIM; 600044; -.
DR   GO; GO:0008083; F:growth factor activity; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   GO; GO:0007275; P:development; TAS.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003978; thrombopoeitin.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR01485; THROMBOPTN.
DR   PROSITE; PS00817;'EPO_TPO; 1.
KW   Cytokine; Glycoprotein; Hormone; Signal; Alternative splicing;
KW   Polymorphism.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    353       THROMBOPOIETIN.
FT   DISULFID     28    172       POTENTIAL.
FT   DISULFID     50    106       POTENTIAL.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    206    206       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    255    255       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    340    340       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    348    348       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    133    136       Missing (in isoform 2).
FT                                /FTId=VSP_001450.
FT   VARSPLIC    160    198       Missing (in isoform 3).
FT                                /FTId=VSP_001451.
FT   VARIANT      14     14       L -> P (IN dbSNP:1042346).
FT                                /FTId=VAR_011795.
FT   VARIANT     116    116       G -> E (IN dbSNP:1126665).
FT                                /FTId=VAR_011796.
FT   CONFLICT     46     46       R -> K (IN REF. 8).
FT   CONFLICT     76     76       M -> MSQ (IN REF. 7).
FT   CONFLICT    113    113       Q -> E (IN REF. 2).
FT   CONFLICT    131    131       T -> P (IN REF. 7).
FT   CONFLICT    277    277       G -> E (IN REF. 8; AAB03393/AAB03394).
FT   CONFLICT    346    346       S -> C (IN REF. 8; AAB03393/AAB03394).
SQ   SEQUENCE   353 AA;  37822 MW;  F0AB5449B72E5526 CRC64;

  Query Match              9.9%;  Score 98.5;  DB 1;  Length 353;
  Best Local Similarity   26.7%;  Pred. No. 0.065;
  Matches   46;  Conservative   21;  Mismatches   82;  Indels   23;  Gaps    6;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDT 71
              :||    | |  |   |||   || ||| : | ::    :  : | |   |   : :|
Db         11 MLLLTARLTLSSP---APP--ACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAV 65

Qy         72 KVNFYAWKRNXSXQQAVEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGL 129
```

```
              :   ||      :| ::   : || | |:  |||      |     |   |   :  |
Db         66 DFSLGEWKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLLLGAL 125

Qy        130 RSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLK 181
              :||      || |     ||        ||        :|  : : ||||::
Db        126 QSL------LGTQ----LPPQG------RTTAHKDPNAIFLSFQHLLRGKVR 161


RESULT 13
TPO_MOUSE
ID   TPO_MOUSE      STANDARD;      PRT;   356 AA.
AC   P40226;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE   (Myeloproliferative leukemia virus oncogene ligand) (C-mpl ligand)
DE   (ML) (Megakaryocyte growth and development factor) (MGDF).
GN   THPO.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Kidney;
RX   MEDLINE=94261207; PubMed=8202158;
RA   Lok S., Kaushansky K., Holly R.D., Kuijper J.L., Lofton-Day C.E.,
RA   Oort P.J., Grant F.J., Heipel M.D., Burkhead S.K., Kramer J.M.,
RA   Bell L.A.N., Sprecher C.A., Blumberg H., Johnson R., Prunkard D.,
RA   Ching A.F.T., Mathewes S.L., Bailey M.C., Forstrom J.W., Buddle M.M.,
RA   Osborn S.G., Evans S.J., Sheppard P.O., Presnell S.R., O'Hara P.J.,
RA   Hagen F.S., Roth G.J., Foster D.C.;
RT   "Cloning and expression of murine thrombopoietin cDNA and stimulation
RT   of platelet production in vivo.";
RL   Nature 369:565-568(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Liver;
RX   MEDLINE=94291201; PubMed=8020099;
RA   Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
RA   Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
RA   Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
RA   Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangelinan M.,
RA   Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
RA   Shutter J., Chute H., Basu R., Selander L., Trollinger D., Sieu L.,
RA   Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA   Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
RA   Pacifici R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
RA   Bosselman R.A.;
RT   "Identification and cloning of a megakaryocyte growth and development
RT   factor that is a ligand for the cytokine receptor Mpl.";
RL   Cell 77:1117-1117(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX   MEDLINE=95152076; PubMed=7849319;
RA   Gurney A.L., Kuang W.-J., Xie M.-H., Malloy B.E., Eaton D.L.,
RA   de Sauvage F.J.;
RT   "Genomic structure, chromosomal localization, and conserved
RT   alternative splice forms of thrombopoietin.";
RL   Blood 85:981-988(1995).
CC   -!- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
CC       AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
CC       CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
CC       MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P40226-1; Sequence=Displayed;
CC       Name=2; Synonyms=TPO-2;
CC         IsoId=P40226-2; Sequence=VSP_001452;
CC   -!- TISSUE SPECIFICITY: FOUND MAINLY IN THE LIVER, KIDNEY AND SKELETAL
CC       MUSCLE.
CC   -!- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC       TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L34169; AAA40436.1; -.
DR   PIR; S45330; S45330.
DR   MGD; MGI:101875; Thpo.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003978; thrombopoeitin.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR01485; THROMBOPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Cytokine; Glycoprotein; Hormone; Signal; Alternative splicing.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    356       THROMBOPOIETIN.
FT   DISULFID    28    172       POTENTIAL.
FT   DISULFID    50    106       POTENTIAL.
FT   CARBOHYD   197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   206    206       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   235    235       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   249    249       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   256    256       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   351    351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   133    136       Missing (in isoform 2).
FT                               /FTId=VSP_001452.
SQ   SEQUENCE   356 AA;  37835 MW;  D346DBAE8392053C CRC64;

  Query Match             9.2%;  Score 91;  DB 1;  Length 356;
  Best Local Similarity   25.1%;  Pred. No. 0.34;
  Matches   44;  Conservative   28;  Mismatches   79;  Indels   24;  Gaps    6;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDT 71
              |||: : | :    | :|    || |:| : | ::    :  : | :   |:  : :|
Db          6 LLLAAMLLAVARLTLSSPVAPACDPRLLNKLLRDSHLLHSRLSQCPDVDPLSIPVLLPAV 65

Qy         72 KVNFYAWKRNXSXQQAVEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGL 129
                 :   ||      :| ::   ::|| | |:  |||          ||  |      | |
Db         66 DFSLGEWKTQTEQSKAQDILGAVSLLLEGVMAARGQL---------EPSCLS-----SLL 111

Qy        130 RSLTTLLR-ALGAQKEAISPPDAASAAPL--RTITADTFRKLFRVYSNFLRGKLK 181
               |:  :|  ||| :  :      :  ||  ||        ||      ||||::
Db        112 GQLSGQVRLLLGALQGLL-----GTQLPLQGRTTAHKDPNALFLSLQQLLRGKVR 161


RESULT 14
TPO_RAT
ID   TPO_RAT        STANDARD;      PRT;   326 AA.
AC   P49745;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Thrombopoietin precursor.
GN   THPO.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=95331639; PubMed=7607561;
RA   Ogami K., Shimada Y., Sohma Y., Akahori H., Kato T., Kawamura K.,
RA   Miyazaki H.;
```

```
RT   "The sequence of a rat cDNA encoding thrombopoietin.";
RL   Gene 158:309-310(1995).
CC   -!- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
CC       AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
CC       CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
CC       MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC       TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D32207; BAA06906.1; -.
DR   PIR; JC4125; JC4125.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003978; thrombopoeitin.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR01485; THROMBOPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Cytokine; Glycoprotein; Hormone; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    326       THROMBOPOIETIN.
FT   DISULFID     28    172       POTENTIAL.
FT   DISULFID     50    106       POTENTIAL.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    206    206       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    235    235       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    249    249       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    256    256       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   326 AA;  34556 MW;  F99D7F77F896FA2C CRC64;

  Query Match             8.8%;  Score 87.5;  DB 1;  Length 326;
  Best Local Similarity   23.0%;  Pred. No. 0.65;
  Matches   42;  Conservative   26;  Mismatches   74;  Indels   41;  Gaps    6;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDT 71
              ||:::| |   |  | :|    || |:| : | ::    :  : | :   |:  : :|
Db          7 LLVAILLLTARL-TLSSPVPPACDPRLLNKLLRDSYLLHSRLSQCPDVNPLSIPVLLPAV 65

Qy         72 KVNFYAWKRNXSXQQAVEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGL 129
                 :  ||      :| ::   ::|| | |:  |||          ||
Db         66 DFSLGEWKTQTEQSKAQDILGAVSLLLEGVMAARGQL---------EP---------- 104

Qy        130 RSLTTLLRALGAQKEAI----------SPPDAASAAPLRTITADTFRKLFRVYSNFLRG 178
               |::||  |  |   :           ||       ||       ||      |||
Db        105 SCLSSLLGQLSGQVRLLLGALQGLLGTQLPPQG------RTTAHKDPSALFLSLQQLLRG 158

Qy        179 KLK 181
               |::
Db        159 KVR 161


RESULT 15
FACG_HUMAN
ID   FACG_HUMAN     STANDARD;      PRT;   622 AA.
AC   O15287;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Fanconi anemia group G protein (FACG protein) (DNA-repair protein
DE   XRCC9).
GN   FANCG OR XRCC9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97404378; PubMed=9256465;
RA   Liu N., Lamerdin J.E., Tucker J.D., Zhou Z.-Q., Walter C.A.,
RA   Albala J.S., Busch D.B., Thompson L.H.;
RT   "The human XRCC9 gene corrects chromosomal instability and mutagen
RT   sensitivities in CHO UV40 cells.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:9232-9237(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99021385; PubMed=9806548;
RA   De Winter J.P., Waisfisz Q., Rooimans M.A., Van Berkel C.G.M.,
RA   Bosnoyan-Collins L., Alon N., Carreau M., Bender O., Demuth I.,
RA   Schindler D., Pronk J.C., Arwert F., Hoehn H., Digweed M.,
RA   Buchwald M., Joenje H.;
RT   "The Fanconi anaemia group G gene FANCG is identical with XRCC9.";
RL   Nat. Genet. 20:281-283(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast, and Foreskin;
RA   Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA   Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA   Phan H., Velasco N., Garnes J., Danganan L., Poundstone P.,
RA   Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
RA   Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
RA   Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A.,
RA   Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
RA   Carrano A.V.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   CHARACTERIZATION.
RX   MEDLINE=99303779; PubMed=10373536;
RA   Garcia-Higuera I., Kuang Y., Naf D., Wasik J., D'Andrea A.D.;
RT   "Fanconi anemia proteins FANCA, FANCC, and FANCG/XRCC9 interact in a
RT   functional nuclear complex.";
RL   Mol. Cell. Biol. 19:4866-4873(1999).
CC   -!- FUNCTION: DNA REPAIR PROTEIN THAT MAY OPERATE IN A POSTREPLICATION
CC       REPAIR OR A CELL CYCLE CHECKPOINT FUNCTION. MAY BE IMPLICATED IN
CC       INTERSTRAND DNA CROSS-LINK REPAIR AND IN THE MAINTENANCE OF NORMAL
CC       CHROMOSOME STABILITY. CANDIDATE TUMOR SUPPRESSOR GENE.
CC   -!- SUBUNIT: Belongs to a multisubunit complex composed of FANCA,
CC       FANCC, FANCF and FANCG proteins. The complex is not found in FA
CC       patients.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (MAJOR) AND CYTOPLASMIC (MINOR).
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS AND THYMUS. FOUND
CC       IN LYMPHOBLASTS.
CC   -!- DISEASE: Defects in FANCG are the cause of one of the 8
CC       complementation groups of Fanconi anemia (FA), an autosomal
CC       recessive disorder characterized by progressive pancytopaenia, a
```

```
CC       diverse assortment of congenital malformations, and a
CC       predisposition to the development of malignancies. At the cellular
CC       level it is associated with hypersensitivity to DNA-damaging
CC       agents, chromosomal instability (increased chromosome breakage),
CC       and defective DNA repair.
CC   -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC       WWW="http://www.infobiogen.fr/services/chromcancer/Genes/FANCGID295.html".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U70310; AAB80802.1; -.
DR   EMBL; AJ007669; CAA07602.1; -.
DR   EMBL; AC004472; AAC07981.1; -.
DR   EMBL; BC000032; AAH00032.1; -.
DR   PIR; T02244; T02244.
DR   Genew; HGNC:3588; FANCG.
DR   MIM; 602956; -.
DR   GO; GO:0003685; F:DNA repair protein; TAS.
DR   GO; GO:0008181; F:tumor suppressor; TAS.
DR   GO; GO:0000075; P:cell cycle checkpoint; TAS.
DR   GO; GO:0006281; P:DNA repair; TAS.
DR   InterPro; IPR001440; TPR.
DR   Pfam; PF00515; TPR; 2.
KW   DNA repair; Nuclear protein.
SQ   SEQUENCE   622 AA;  68553 MW;  4BC7475472AC3C84 CRC64;

  Query Match             8.7%;  Score 86.5;  DB 1;  Length 622;
  Best Local Similarity   25.8%;  Pred. No. 1.7;
  Matches   49;  Conservative   23;  Mismatches   83;  Indels   35;  Gaps    6;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC--SLNENITVP 69
              || ||  ||  :|||     : |:       :| |  |:  |   |:    ||   :
Db         53 LLHSLQGLPAAVPVLPLELTVTCNF-----IILRASLAQGFTEDQAQDIQRSLERVLETQ 107

Qy         70 DTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQAL---LVNSSQPWEPLQLHVDKAV 126
              :              ||   : |||  | ::|||   |   |:::      ||  :  :
Db        108 E--------------QQGPRLEQGLRELWDSVLRASCLLPELLSALHRLVGLQAALWLSA 153

Qy        127 SGLRSLTTLLRAL-----GAQKEAI------SPPDAASAAPLRTITADTFRKLFRVYSNF 175
              |  |  ||  |     || |: :      |||     |||     |   : :       :
Db        154 DRLGDLALLLETLNGSQSGASKDLLLLLKTWSPPAEELDAPLTLQDAQGLKDVLLTAFAY 213

Qy        176 LRGKLKLYTG 185
               :|  :| ||
Db        214 RQGLQELITG 223


Search completed: September  8, 2003, 15:01:17
Job time : 18 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 14:57:36 ; Search time 20 Seconds
(without alignments)
928.028 Million cell updates/sec

Title: US-09-813-775C-34
Perfect score: 991
Sequence: 1 MGVHECPAWLWLLLSLLSLP.....NFLRGKLKLYTGEACRTGDR 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	973	98.2	193	1	ZUHU	erythropoietin pre
2	881.5	89.0	192	1	JQ0173	erythropoietin pre
3	879.5	88.7	192	1	I84613	erythropoietin pre
4	779.5	78.7	188	1	I46083	erythropoietin pre
5	762.5	76.9	192	1	S28148	erythropoietin pre
6	752.5	75.9	194	1	I46401	erythropoietin pre
7	748.5	75.5	190	2	I46578	erythropoietin - p
8	738.5	74.5	192	1	A24902	erythropoietin pre
9	729	73.6	195	2	JC7699	erythropoietin - r
10	699.5	70.6	175	2	I46199	erythropoietin - d
11	109	11.0	353	2	G02729	thrombopoietin - h
12	98.5	9.9	353	2	I80105	thrombopoietin pre
13	95.5	9.6	286	2	A55530	megakaryocyte grow
14	91	9.2	356	2	S45330	thrombopoietin - m
15	87.5	8.8	326	2	JC4125	thrombopoietin pre
16	87.5	8.8	346	2	AE0959	Solute binding rec
17	86.5	8.7	622	2	T02244	probable DNA repai
18	85	8.6	235	2	JC4227	thrombopoietin pre
19	84	8.5	339	2	A83274	UDP-N-acetylpyruvo
20	81.5	8.2	480	2	S56639	ribosomal protein
21	81	8.2	323	2	AB0323	ribonucleoside-dip
22	80	8.1	353	2	AG0138	quinolinate synthe
23	80	8.1	1336	2	T18288	ABC transport prot
24	80	8.1	1980	2	S54307	myosin heavy chain
25	79.5	8.0	475	2	D84064	succinate-semialde
26	79	8.0	743	2	D75590	methyl-accepting c
27	79	8.0	1089	2	S53978	PSE1 protein - yea
28	78.5	7.9	296	2	AI0443	probable 2-hydroxy
29	78.5	7.9	348	2	T35450	ABC transporter AT
30	78.5	7.9	2201	2	AH0095	probable sideropho
31	78	7.9	384	2	AI2962	cellulose synthesi
32	78	7.9	389	2	E98320	hypothetical prote
33	78	7.9	451	2	S75569	hypothetical prote
34	78	7.9	1189	2	I39711	celE protein - Agr
35	77.5	7.8	242	2	AD1928	hypothetical prote
36	77.5	7.8	567	2	T08405	hypothetical prote
37	77.5	7.8	774	2	F96639	protein T1F9.8 [im
38	77	7.8	416	2	B84276	glutamate dehydrog
39	76.5	7.7	544	2	S37039	groEL protein - Ba
40	76.5	7.7	1638	2	T30313	chemotaxis protein
41	76.5	7.7	2472	2	E83594	still frameshift p
42	76	7.7	552	2	S39765	chaperonin 60 - Co
43	76	7.7	554	2	T27878	hypothetical prote
44	75	7.6	263	2	B75361	WD-repeat family p
45	74.5	7.5	455	2	AG2919	conserved hypothet

ALIGNMENTS

RESULT 1
ZUHU
erythropoietin precursor [validated] - human
C;Species: Homo sapiens (man)
C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C;Accession: A01855; A24744; A25384; A22210; S56178
R;Jacobs, K.; Shoemaker, C.; Rudersdorf, R.; Neill, S.D.; Kaufman, R.J.; Mufson, A.;
Nature 313, 806-810, 1985
A;Title: Isolation and characterization of genomic and cDNA clones of human erythropo
A;Reference number: A01855; MUID:85137899; PMID:3838366
A;Accession: A01855
A;Molecule type: mRNA; DNA
A;Residues: 1-193 <JAC>
A;Cross-references: GB:X02157; GB:X02158
R;Lin, F.K.; Suggs, S.; Lin, C.H.; Browne, J.K.; Smalling, R.; Egrie, J.C.; Chen, K.K
Proc. Natl. Acad. Sci. U.S.A. 82, 7580-7584, 1985
A;Title: Cloning and expression of the human erythropoietin gene.
A;Reference number: A24744; MUID:86067948; PMID:3865178
A;Accession: A24744
A;Molecule type: DNA
A;Residues: 1-193 <LIN>
A;Cross-references: GB:M11319; NID:g182197; PIDN:AAA52400.1; PID:g182198
R;Lai, P.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser, E.
J. Biol. Chem. 261, 3116-3121, 1986
A;Title: Structural characterization of human erythropoietin.
A;Reference number: A25384; MUID:86140080; PMID:3949763
A;Accession: A25384
A;Molecule type: protein
A;Residues: 28-86,'Q',87-193 <LAI>
A;Experimental source: urine
A;Note: forms without the carboxyl-terminal residue and the four carboxyl-terminal re
R;Yanagawa, S.; Hirade, K.; Ohnota, H.; Sasaki, R.; Chiba, H.; Ueda, M.; Goto, M.
J. Biol. Chem. 259, 2707-2710, 1984
A;Title: Isolation of human erythropoietin with monoclonal antibodies.
A;Reference number: A22210; MUID:84135751; PMID:6698989
A;Accession: A22210
A;Molecule type: protein
A;Residues: 28-29,'X',31-33,'L',35-50,'X',52-53,'D',55,'G',57 <YAN>
R;Matsumoto, S.; Ikura, K.; Ueda, M.; Sasaki, R.
Plant Mol. Biol. 27, 1163-1172, 1995
A;Title: Characterization of a human glycoprotein (erythropoietin) produced in cultur
A;Reference number: S56178; MUID:95284365; PMID:7766897
A;Accession: S56178
A;Molecule type: protein
A;Residues: 28-33,'X',35-37 <MTS>
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C;Genetics:
A;Gene: GDB:EPO
A;Cross-references: GDB:119110; OMIM:133170
A;Map position: 7q21.3-7q22.1
A;Introns: 5/1; 53/3; 82/3; 142/3
C;Function:

```
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-193/Product: erythropoietin #status experimental <MAT>
F;34-188,56-60/Disulfide bonds: #status experimental
F;51,65,110/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;153/Binding site: carbohydrate (Ser) (covalent) #status experimental

  Query Match             98.2%;  Score 973;  DB 1;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 1.4e-86;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 2
JQ0173
erythropoietin precursor - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 07-Sep-1990 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: JQ0173
R;Lin, F.K.; Lin, C.H.; Lai, P.H.; Browne, J.K.; Egrie, J.C.; Smalling, R.; Fox, G.M.; C
Gene 44, 201-209, 1986
A;Title: Monkey erythropoietin gene: cloning, expression and comparison with the human e
A;Reference number: JQ0173; MUID:87055236; PMID:2877922
A;Accession: JQ0173
A;Molecule type: mRNA
A;Residues: 1-192 <LIN>
A;Cross-references: GB:M18189; GB:M15818; GB:M15819; GB:M18188; NID:g342093; PIDN:AAA368
A;Experimental source: kidney
C;Comment: This protein is the principal hormone involved in the regulation of erythrocy
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver c
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-192/Product: erythropoietin #status predicted <MAT>
F;34-187,56-60/Disulfide bonds: #status predicted
F;51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;152/Binding site: carbohydrate (Ser) (covalent) #status predicted

  Query Match             89.0%;  Score 881.5;  DB 1;  Length 192;
  Best Local Similarity   89.6%;  Pred. No. 9.6e-78;
  Matches  173;  Conservative    8;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||:|||||||| |||||||||||||||||||||||:| ||:| |
Db          1 MGVHECPAWLWLLLSLVSLPLGLPVPGAPPRLICDSRVLERYLLEAKEAENVTMGCSESC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||    |||||||||||||||||||||:| |||||:|||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQAVLANSSQPFEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              |:|||:|||||:|||||||||| |||| |||||||||||||||| ||||||||||||
Db        121 HMDKAISGLRSITTLLRALGAQ-EAISLPDAASAAPLRTITADTFCKLFRVYSNFLRGKL 179

Qy        181 KLYTGEACRTGDR 193
              ||||||||| |||
Db        180 KLYTGEACRRGDR 192


RESULT 3
I84613
erythropoietin precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 02-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: I84613
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence hom
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I84613
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-192 <RES>
A;Cross-references: GB:L10609; NID:g342095; PIDN:AAA36842.1; PID:g342096
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-192/Product: erythropoietin #status predicted <MAT>
F;34-187,56-60/Disulfide bonds: #status predicted
F;51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;152/Binding site: carbohydrate (Ser) (covalent) #status predicted

  Query Match             88.7%;  Score 879.5;  DB 1;  Length 192;
  Best Local Similarity   89.1%;  Pred. No. 1.5e-77;
  Matches  172;  Conservative    9;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||:|||||||| |||||:|||||||||||||||||:| ||:| |
Db          1 MGVHECPAWLWLLLSLVSLPLGLPVPGAPPRLVCDSRVLERYLLEAKEAENVTMGCSESC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||    |||||||||||||||||||||:| |||||:|||||
Db         61 SLNENITVPDTKVNFYAWKRIEVGQQAVEVWQGLALLSEAVLRGQAVLANSSQPFEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              |:|||:|||||:|||||||||| |||| |||||||||||||||| ||||||||||||
Db        121 HMDKAISGLRSITTLLRALGAQ-EAISLPDAASAAPLRTITADTFCKLFRVYSNFLRGKL 179

Qy        181 KLYTGEACRTGDR 193
              ||||||||| |||
Db        180 KLYTGEACRRGDR 192


RESULT 4
I46083
erythropoietin precursor - cat (fragment)
C;Species: Felis silvestris catus (domestic cat)
C;Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: I46083
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence hom
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I46083
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-188 <WEN>
A;Cross-references: GB:L10606; NID:g163820; PIDN:AAA30807.1; PID:g163821
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-22/Domain: signal sequence (fragment) #status predicted <SIG>
```

```
F;23-188/Product: erythropoietin #status predicted <MAT>
F;29-183,51-55/Disulfide bonds: #status predicted
F;46,60,105/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;148/Binding site: carbohydrate (Ser) (covalent) #status predicted

  Query Match             78.7%;  Score 779.5;  DB 1;  Length 188;
  Best Local Similarity   83.1%;  Pred. No. 6.8e-68;
  Matches  157;  Conservative    8;  Mismatches   23;  Indels    1;  Gaps    1;

Qy          5 ECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNE 64
              |||| | |||||| |||||||||||||||||||||||||:|||:||||:| |||| || :|
Db          1 ECPA-LLLLLSLLLLPLGLPVLGAPPRLICDSRVLERYILEAREAENVTMGCAEGCSFSE 59

Qy         65 NITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDK 124
              |||||||||||| |||    |||||||||||||||||:||||||| ||||| | |||||||
Db         60 NITVPDTKVNFYTWKRMDVGQQAVEVWQGLALLSEAILRGQALLANSSQPSETLQLHVDK 119

Qy        125 AVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYT 184
              ||| |||||:|||||||||||| | |:| ||||||| | ||  ||||:||||||||| |||
Db        120 AVSSLRSLTSLLRALGAQKEATSLPEATSAAPLRTFTVDTLCKLFRIYSNFLRGKLTLYT 179

Qy        185 GEACRTGDR 193
              ||||| |||
Db        180 GEACRRGDR 188


RESULT 5
S28148
erythropoietin precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 15-Nov-1996 #text_change 21-Jul-2000
C;Accession: S28148; I62743
R;Nagao, M.; Suga, H.; Okano, M.; Masuda, S.; Narita, H.; Ikura, K.; Sasaki, R.
Biochim. Biophys. Acta 1171, 99-102, 1992
A;Title: Nucleotide sequence of rat erythropoietin.
A;Reference number: S28148; MUID:93042015; PMID:1420369
A;Accession: S28148
A;Molecule type: mRNA
A;Residues: 1-192 <NAG>
A;Cross-references: GB:D10763; NID:g220735; PIDN:BAA01593.1; PID:g220736
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence homolo
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I62743
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 4-192 <RES>
A;Cross-references: GB:L10608; NID:g204060; PIDN:AAA41126.1; PID:g204061
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver c
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-192/Product: erythropoietin #status predicted <MAT>
F;33-187,55-165/Disulfide bonds: #status predicted
F;50,64,109/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             76.9%;  Score 762.5;  DB 1;  Length 192;
  Best Local Similarity   79.8%;  Pred. No. 3.1e-66;
  Matches  154;  Conservative   13;  Mismatches   25;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||| | |  | |||||| :||||||| |||||||||||||||:||||||||:| ||||
Db          1 MGVPERPT-LLLLLSLLLIPLGLPVLCAPPRLICDSRVLERYILEAKEAENVTMGCAEGP 59

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
               |:|||||||||||||||    :|||||||||:|||||:|: |||  ||||| | |||
Db         60 RLSENITVPDTKVNFYAWKRMKVEEQAVEVWQGLSLLSEAILQAQALQANSSQPPESLQL 119

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              |:|||:|||||||:||| |||||| :|||||  ||||||:||||| ||||||||||||||
Db        120 HIDKAISGLRSLTSLLRVLGAQKELMSPPDATQAAPLRTLTADTFCKLFRVYSNFLRGKL 179

Qy        181 KLYTGEACRTGDR 193
              ||||||||| |||
Db        180 KLYTGEACRRGDR 192


RESULT 6
I46401
erythropoietin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: I46401; I47077
R;Fu, P.; Evans, B.; Lim, G.B.; Moritz, K.; Wintour, E.M.
Mol. Cell. Endocrinol. 93, 107-116, 1993
A;Title: The sheep erythropoietin gene: molecular cloning and effect of hemorrhage on
A;Reference number: I46401; MUID:93351736; PMID:8349021
A;Accession: I46401
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-194 <FUX>
A;Cross-references: EMBL:Z24681; NID:g395049; PIDN:CAA80848.1; PID:g395050
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence hom
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I47077
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 4-15,'L',17-107,'P',109-194 <WEN>
A;Cross-references: GB:L10610; NID:g165876; PIDN:AAA31518.1; PID:g165877
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-194/Product: erythropoietin #status predicted <MAT>
F;34-189,56-60/Disulfide bonds: #status predicted
F;51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;154/Binding site: carbohydrate (Ser) (covalent) #status predicted

  Query Match             75.9%;  Score 752.5;  DB 1;  Length 194;
  Best Local Similarity   77.8%;  Pred. No. 2.9e-65;
  Matches  151;  Conservative   10;  Mismatches   32;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||  :|   | |||| |  |||||||||||||||||||||:|||:|||| | |||| |
Db          1 MGARDCTPLLLLLLSFLLFPLGLPVLGAPPRLICDSRVLERYILEAREAENATMGCAEGC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              | :|||||||||||||||||    |||:||||||||||||: |||||| |:||| | |:|
Db         61 SFSENITVPDTKVNFYAWKRMEVQQQALEVWQGLALLSEAIFRGQALLANASQPCEALRL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAA-SAAPLRTITADTFRKLFRVYSNFLRGK 179
              |||||||||||||:|||||||||||||  |||  ||||||  | |   ||||:||||||||
Db        121 HVDKAVSGLRSLTSLLRALGAQKEAIPLPDATPSAAPLRIFTVDALSKLFRIYSNFLRGK 180

Qy        180 LKLYTGEACRTGDR 193
              | |||||||| |||
Db        181 LTLYTGEACRRGDR 194


RESULT 7
I46578
erythropoietin - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C;Accession: I46578
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence hom
A;Reference number: I46083; MUID:93372347; PMID:8364201
```

```
A;Accession: I46578
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-190 <WEN>
A;Cross-references: GB:L10607; NID:g164445; PIDN:AAA31029.1; PID:g164446
C;Superfamily: erythropoietin

  Query Match             75.5%;  Score 748.5;  DB 2;  Length 190;
  Best Local Similarity   81.2%;  Pred. No. 6.8e-65;
  Matches  155;  Conservative    7;  Mismatches   26;  Indels    3;  Gaps    2;

Qy          5 ECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNE 64
              |||| | |||||| |||||||||||||||||||||||||:||||| || | |||| || :|
Db          1 ECPARL-LLLSLLLLPLGLPVLGAPPRLICDSRVLERYILEAKEGENATMGCAESCSFSE 59

Qy         65 NITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDK 124
              ||||||||||||||||    |||:||||||||||||:|:||||| ||||| | |||||||
Db         60 NITVPDTKVNFYAWKRMEVQQQAMEVWQGLALLSEAILQGQALLANSSQPSEALQLHVDK 119

Qy        125 AVSGLRSLTTLLRALGAQKEAISPPDA--ASAAPLRTITADTFRKLFRVYSNFLRGKLKL 182
              |||||||||:|||||||||||||  |||  :|| ||||   ||  |||| |||||||||  |
Db        120 AVSGLRSLTSLLRALGAQKEAIPLPDASPSSATPLRTFAVDTLCKLFRNYSNFLRGKLTL 179

Qy        183 YTGEACRTGDR 193
              |||||||  ||
Db        180 YTGEACRRRDR 190


RESULT 8
A24902
erythropoietin precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 25-Oct-1987 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: A24902; A24901
R;Shoemaker, C.B.; Mitsock, L.D.
Mol. Cell. Biol. 6, 849-858, 1986
A;Title: Murine erythropoietin gene: cloning, expression, and human gene homology.
A;Reference number: A24902; MUID:87039105; PMID:3773894
A;Accession: A24902
A;Molecule type: DNA
A;Residues: 1-192 <SHO>
A;Note: the authors translated the codon TTA for residue 12 as Phe, TTA for residue 43 a
R;McDonald, J.D.; Lin, F.K.; Goldwasser, E.
Mol. Cell. Biol. 6, 842-848, 1986
A;Title: Cloning, sequencing, and evolutionary analysis of the mouse erythropoietin gene
A;Reference number: A24901; MUID:87039104; PMID:3022133
A;Accession: A24901
A;Molecule type: DNA
A;Residues: 1-67,'P',69-192 <MCD>
A;Cross-references: GB:M12930; NID:g193086; PIDN:AAA37570.1; PID:g387152
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver d
C;Genetics:
A;Introns: 5/1; 52/3; 81/3; 141/3
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-192/Product: erythropoietin #status predicted <MAT>
F;33-187,55-165/Disulfide bonds: #status predicted
F;50,64,109/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             74.5%;  Score 738.5;  DB 1;  Length 192;
  Best Local Similarity   76.7%;  Pred. No. 6.4e-64;
  Matches  148;  Conservative   15;  Mismatches   29;  Indels    1;  Gaps    1;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||| | |  | |||||| :||||||| |||||||||||||||:||||||||:| ||||
Db          1 MGVPERPT-LLLLLSLLLIPLGLPVLCAPPRLICDSRVLERYILEAKEAENVTMGCAEGP 59

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
               |:||||| ||||||||||    :|||:||||||:|||||:|: |||| ||||| | |||
Db         60 RLSENITVADTKVNFYAWKRMEVEEQAIEVWQGLSLLSEAILQAQALLANSSQPPETLQL 119

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              |:|||:|||||||:||| |||||| :|||    |||||:| ||| ||||||:|||||||
Db        120 HIDKAISGLRSLTSLLRVLGAQKELMSPPDTTPPAPLRTLTVDTFCKLFRVYANFLRGKL 179

Qy        181 KLYTGEACRTGDR 193
              |||||| || |||
Db        180 KLYTGEVCRRGDR 192


RESULT 9
JC7699
erythropoietin - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C;Accession: JC7699
R;Vilalta, A.; Wu, D.; Margalith, M.; Hobart, P.
Biochem. Biophys. Res. Commun. 284, 823-827, 2001
A;Title: Rabbit EPO gene and cDNA: Expression of rabbit EPO after intramuscular injec
A;Reference number: JC7699; MUID:21290682; PMID:11396976
A;Contents: Kidney
A;Accession: JC7699
A;Molecule type: DNA
A;Residues: 1-195 <VIL>
A;Cross-references: GB:AF290943
C;Comment: This protein, a heavily glycosylated 34K protein produced in the fetal liv
cytes.
C;Genetics:
A;Gene: epo
C;Superfamily: erythropoietin
C;Keywords: glycoprotein; kidney

  Query Match             73.6%;  Score 729;  DB 2;  Length 195;
  Best Local Similarity   77.9%;  Pred. No. 5.4e-63;
  Matches  152;  Conservative   14;  Mismatches   27;  Indels    2;  Gaps    2;

Qy          1 MGVHECPAWLWL-LLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEH 59
              |||    | | | || || | ||||||||| ||||||||||||:|||||||:| ||||
Db          1 MGVRGRLALLPLALLCLLVLALGLPVLGAPARLICDSRVLERYILEAKEAENVTMGCAEG 60

Qy         60 CSLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQ 119
              ||| |||||||||||||: ||:: : : |||||||||||||||:|| |||| |||| | ||
Db         61 CSLGENITVPDTKVNFHHWKKSEAGRHAVEVWQGLALLSEAMLRSQALLANSSQLPETLQ 120

Qy        120 LHVDKAVSGLRSLTTLLRALGAQKEAISPPDAA-SAAPLRTITADTFRKLFRVYSNFLRG 178
              :|||||||||||||:|||||| ||||:|||:|| |||||||: ||| |||||||||||||
Db        121 VHVDKAVSGLRSLTSLLRALGVQKEAVSPPEAASSAAPLRTVAADTLCKLFRIYSNFLRG 180

Qy        179 KLKLYTGEACRTGDR 193
              |||||||||||  |||
Db        181 KLKLYTGEACRRGDR 195


RESULT 10
I46199
erythropoietin - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 23-Jul-1999
C;Accession: I46199
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence hom
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I46199
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-175 <WEN>
A;Cross-references: GB:L13027; NID:g290087; PIDN:AAA30842.1; PID:g552347
C;Superfamily: erythropoietin

  Query Match             70.6%;  Score 699.5;  DB 2;  Length 175;
  Best Local Similarity   80.1%;  Pred. No. 3.4e-60;
  Matches  141;  Conservative   12;  Mismatches   22;  Indels    1;  Gaps    1;
```

```
Qy          5 ECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNE 64
              |||| | |||||| |||||||||||||||||||||||||:|||:||||:| |||: || :|
Db          1 ECPA-LLLLLSLLLLPLGLPVLGAPPRLICDSRVLERYILEAREAENVTMGCAQGCSFSE 59

Qy         65 NITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDK 124
              |||||||||||| |||    |||:||||||||||||:||||||| |:||| |  ||||||
Db         60 NITVPDTKVNFYTWKRMDVGQQALEVWQGLALLSEAILRGQALLANASQPSETPQLHVDK 119

Qy        125 AVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||| |||||:||||||||||:| |: || ||||| | ||  |||||:||||||||
Db        120 AVSSLRSLTSLLRALGAQKEAMSLPEEASPAPLRTFTVDTLCKLFRIYSNFLRGKL 175


RESULT 11
G02729
thrombopoietin - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 05-Nov-1999
C;Accession: G02729
R;Im, S.
submitted to the EMBL Data Library, May 1996
A;Reference number: H01637
A;Accession: G02729
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-353 <IMX>
A;Cross-references: EMBL:U59493; NID:g1401245; PIDN:AAB03392.1; PID:g1401246
C;Genetics:
A;Gene: hTPO

  Query Match             11.0%;  Score 109;  DB 2;  Length 353;
  Best Local Similarity   26.2%;  Pred. No. 0.0097;
  Matches   45;  Conservative   23;  Mismatches   86;  Indels   18;  Gaps    4;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDT 71
              ||| :: ||     | :|    || ||| : | ::    :  : | |   |   : :|
Db          6 LLLVVMLLPTARLTLSSPAPPACDLRVLSKLLRDSHVLHSKLSQCPEVHPLPTPVLLPAV 65

Qy         72 KVNFYAWKRNXSXQQAVEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGL 129
                :  ||      :| ::   : || | |:  |||      |     |  |   :  |
Db         66 DFSLGEWKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSEQVRLLLGAL 125

Qy        130 RSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLK 181
              :||      || |     ||        ||        :|  : : ||||::
Db        126 QSL------LGTQ----LPPQG------RTTAHKDPNAIFLSFQHLLRGKVR 161


RESULT 12
I80105
thrombopoietin precursor - human
N;Alternate names: c-MPL ligand; megakaryocyte growth and development factor precursor
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 20-Jun-2000
C;Accession: I59281; I80105; S45331; S48740; I38672; I52610
R;Foster, D.C.; Sprecher, C.A.; Grant, F.J.; Kramer, J.M.; Kuijper, J.L.; Holly, R.D.; W
Proc. Natl. Acad. Sci. U.S.A. 91, 13023-13027, 1994
A;Title: Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosoma
A;Reference number: I59281; MUID:95108091; PMID:7809166
A;Accession: I59281
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-353 <RE2>
A;Cross-references: GB:L36051; NID:g533214; PIDN:AAC37568.1; PID:g533215
A;Accession: I80105
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-353 <RES>
A;Cross-references: GB:L36052; NID:g533216; PIDN:AAC37566.1; PID:g533217
R;de Sauvage, F.J.; Hass, P.E.; Spencer, S.D.; Malloy, B.E.; Gurney, A.L.; Spencer, S.A.
D.V.; Eaton, D.L.
Nature 369, 533-538, 1994
A;Title: Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl ligand.
A;Reference number: S45331; MUID:94261202; PMID:8202154
A;Accession: S45331
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-353 <SAU>
A;Cross-references: GB:L33410; NID:g506826; PIDN:AAA59857.1; PID:g506827
R;Sohma, Y.; Akahori, H.; Seki, N.; Hori, T.; Ogami, K.; Kato, T.; Shimada, Y.; Kawam
FEBS Lett. 353, 57-61, 1994
A;Title: Molecular cloning and chromosomal localization of the human thrombopoietin g
A;Reference number: S48740; MUID:95010765; PMID:7926023
A;Accession: S48740
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-353 <SOH>
A;Cross-references: GB:D32046; NID:g577319; PIDN:BAA06807.1; PID:g577320
R;Bartley, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.
Cell 77, 1117-1124, 1994
A;Title: Identification and cloning of a megakaryocyte growth and development factor
A;Reference number: A54463; MUID:94291201; PMID:8020099
A;Accession: I38672
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-112,'E',114-353 <RE3>
A;Cross-references: EMBL:U11025; NID:g511223; PIDN:AAA50553.1; PID:g558078
R;Gurney, A.L.; Kuang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A;Title: Genomic structure, chromosomal localization, and conserved alternative splic
A;Reference number: I52610; MUID:95152076; PMID:7849319
A;Accession: I52610
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-353 <RE4>
A;Cross-references: GB:S76771; NID:g914225; PIDN:AAB33390.1; PID:g914226
C;Genetics:
A;Gene: GDB:THPO; MGDF
A;Cross-references: GDB:374007; OMIM:600044
A;Map position: 3q26.3-3q27
A;Introns: 5/1; 47/3; 76/3; 132/3
C;Keywords: alternative splicing; cytokine; glycoprotein

  Query Match             9.9%;  Score 98.5;  DB 2;  Length 353;
  Best Local Similarity   26.7%;  Pred. No. 0.1;
  Matches   46;  Conservative   21;  Mismatches   82;  Indels   23;  Gaps    6;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDT 71
              :||    | |  |   |||   || ||| : | ::    :  : | |   |   : :|
Db         11 MLLLTARLTLSSP---APP--ACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAV 65

Qy         72 KVNFYAWKRNXSXQQAVEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGL 129
                :  ||      :| ::   : || | |:  |||      |     |  |   :  |
Db         66 DFSLGEWKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLLLGAL 125

Qy        130 RSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLK 181
              :||      || |     ||        ||        :|  : : ||||::
Db        126 QSL------LGTQ----LPPQG------RTTAHKDPNAIFLSFQHLLRGKVR 161


RESULT 13
A55530
megakaryocyte growth and development factor, long form - human
N;Alternate names: MPL ligand, long form
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
C;Accession: A55530
R;Chang, M.; McNinch, J.; Basu, R.; Shutter, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs
J. Biol. Chem. 270, 511-514, 1995
A;Title: Cloning and characterization of the human megakoryocyte growth and developme
A;Reference number: A55530; MUID:95122483; PMID:7822271
A;Accession: A55530
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-286 <CHA>
A;Cross-references: GB:U17071
```

```
C;Genetics:
A;Gene: MGDF
A;Map position: 3q26.3
C;Keywords: alternative splicing; cytokine

  Query Match             9.6%;  Score 95.5;  DB 2;  Length 286;
  Best Local Similarity   27.1%;  Pred. No. 0.15;
  Matches   46;  Conservative   19;  Mismatches   82;  Indels   23;  Gaps    6;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDT 71
              :||    | |  |   |||   || ||| : | ::    :  : | |   |   : :|
Db         11 MLLLTARLTLSSP---APP--ACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAV 65

Qy         72 KVNFYAWKRNXSXQQAVEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGL 129
                :   ||      :| ::   : || | |:  |||      |     |   |   :  |
Db         66 DFSLGEWKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLLLGAL 125

Qy        130 RSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGK 179
              :||      || |     ||        ||        :|  : : ||||
Db        126 QSL------LGTQ----LPPQG------RTTAHKDPNAIFLSFQHLLRGK 159


RESULT 14
S45330
thrombopoietin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Dec-1994 #sequence_revision 17-Nov-1995 #text_change 08-Oct-1999
C;Accession: S45330
R;Lok, S.; Kaushansky, K.; Holly, R.D.; Kuijper, J.L.; Lofton-Day, C.E.; Oort, P.J.; Gra
unkard, D.; Ching, A.F.T.; Mathewes, S.L.; Bailey, M.C.; Forstrom, J.W.; Buddle, M.M.; C
Nature 369, 565-568, 1994
A;Title: Cloning and expression of murine thrombopoietin cDNA and stimulation of platele
A;Reference number: S45330; MUID:94261207; PMID:8202158
A;Accession: S45330
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-356 <LOK>
A;Cross-references: GB:L34169; NID:g508540; PIDN:AAA40436.1; PID:g508541

  Query Match             9.2%;  Score 91;  DB 2;  Length 356;
  Best Local Similarity   25.1%;  Pred. No. 0.54;
  Matches   44;  Conservative   28;  Mismatches   79;  Indels   24;  Gaps    6;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDT 71
              |||: : | :    | :|    || |:| : | ::    :  : | :   |:  : :|
Db          6 LLLAAMLLAVARLTLSSPVAPACDPRLLNKLLRDSHLLHSRLSQCPDVDPLSIPVLLPAV 65

Qy         72 KVNFYAWKRNXSXQQAVEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGL 129
                :   ||      :| ::   ::|| | |:  |||          ||  |      | |
Db         66 DFSLGEWKTQTEQSKAQDILGAVSLLLEGVMAARGQL---------EPSCLS-----SLL 111

Qy        130 RSLTTLLR-ALGAQKEAISPPDAASAAPL--RTITADTFRKLFRVYSNFLRGKLK 181
                 |:  :|  ||| :  :      :  ||  ||        ||      ||||::
Db        112 GQLSGQVRLLLGALQGLL-----GTQLPLQGRTTAHKDPNALFLSLQQLLRGKVR 161


RESULT 15
JC4125
thrombopoietin precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 20-Jun-2000
C;Accession: JC4125
R;Ogami, K.; Shimada, Y.; Sohma, Y.; Akahori, H.; Kato, T.; Kawamura, K.; Miyazaki, H.
Gene 158, 309-310, 1995
A;Title: The sequence of a rat cDNA encoding thrombopoietin.
A;Reference number: JC4125; MUID:95331639; PMID:7607561
A;Accession: JC4125
A;Molecule type: mRNA
A;Residues: 1-326 <OGA>
A;Cross-references: DDBJ:D32207; NID:g971276; PIDN:BAA06906.1; PID:g971277
A;Experimental source: liver
C;Comment: This protein is a hematopoietic factor which shows megakaryocyte colony-stimu
C;Genetics:
A;Introns: 5/2; 132/3
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-326/Product: thrombopoietin #status predicted <MAT>

  Query Match             8.8%;  Score 87.5;  DB 2;  Length 326;
  Best Local Similarity   23.0%;  Pred. No. 1.1;
  Matches   42;  Conservative   26;  Mismatches   74;  Indels   41;  Gaps    6;

Qy         12 LLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDT 71
              ||:::| |   |  | :|    || |:| : | ::    :  : | :   |:  : :|
Db          7 LLVAILLLTARL-TLSSPVPPACDPRLLNKLLRDSYLLHSRLSQCPDVNPLSIPVLLPAV 65

Qy         72 KVNFYAWKRNXSXQQAVEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGL 129
                :   ||      :| ::   ::|| | |:  |||          ||
Db         66 DFSLGEWKTQTEQSKAQDILGAVSLLLEGVMAARGQL---------EP------------ 104

Qy        130 RSLTTLLRALGAQKEAI----------SPPDAASAAPLRTITADTFRKLFRVYSNFLRG 178
              |::||  |  |  :            ||        ||       ||      |||
Db        105 SCLSSLLGQLSGQVRLLLGALQGLLGTQLPPQG------RTTAHKDPSALFLSLQQLLRG 158

Qy        179 KLK 181
              |::
Db        159 KVR 161


Search completed: September  8, 2003, 15:02:39
Job time : 22 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 15:00:51 ; Search time 13.2862 Seconds
(without alignments)
614.621 Million cell updates/sec

Title: US-09-813-775C-34
Perfect score: 991
Sequence: 1 MGVHECPAWLWLLLSLLSLP.....NFLRGKLKLYTGEACRTGDR 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	987	99.6	193	4	US-09-552-265B-34	Sequence 34, Appl
2	984	99.3	193	4	US-09-552-265B-36	Sequence 36, Appl
3	983	99.2	193	4	US-09-552-265B-35	Sequence 35, Appl
4	980	98.9	193	4	US-09-552-265B-37	Sequence 37, Appl
5	973	98.2	193	1	US-07-903-220-1	Sequence 1, Appli
6	973	98.2	193	2	US-08-883-795A-34	Sequence 34, Appl
7	973	98.2	193	4	US-09-552-265B-2	Sequence 2, Appli
8	973	98.2	193	4	US-09-552-265B-4	Sequence 4, Appli
9	970	97.9	193	4	US-09-552-265B-38	Sequence 38, Appl
10	970	97.9	193	4	US-09-552-265B-40	Sequence 40, Appl
11	969	97.8	193	4	US-09-552-265B-5	Sequence 5, Appli
12	966	97.5	193	4	US-09-552-265B-39	Sequence 39, Appl
13	966	97.5	193	4	US-09-552-265B-41	Sequence 41, Appl
14	965	97.4	193	4	US-09-552-265B-42	Sequence 42, Appl
15	964	97.3	193	4	US-09-552-265B-44	Sequence 44, Appl
16	961	97.0	193	4	US-09-552-265B-43	Sequence 43, Appl
17	960	96.9	193	4	US-09-552-265B-45	Sequence 45, Appl
18	958	96.7	193	4	US-09-552-265B-46	Sequence 46, Appl
19	957	96.6	193	4	US-09-552-265B-48	Sequence 48, Appl
20	954	96.3	193	4	US-09-552-265B-47	Sequence 47, Appl
21	953	96.2	193	4	US-09-552-265B-49	Sequence 49, Appl
22	839	84.7	166	4	US-09-552-265B-18	Sequence 18, Appl
23	836	84.4	166	4	US-09-552-265B-20	Sequence 20, Appl
24	835	84.3	166	4	US-09-552-265B-19	Sequence 19, Appl
25	832	84.0	166	4	US-09-552-265B-21	Sequence 21, Appl
26	828	83.6	412	4	US-09-366-009-34	Sequence 34, Appl
27	828	83.6	412	4	US-08-809-156B-34	Sequence 34, Appl
28	827	83.5	166	5	PCT-US94-04361-45	Sequence 45, Appl
29	825	83.2	166	1	US-08-318-193-70	Sequence 70, Appl
30	825	83.2	166	4	US-09-604-871-2	Sequence 2, Appli
31	825	83.2	166	4	US-09-604-938-2	Sequence 2, Appli
32	825	83.2	166	5	PCT-US94-04361-37	Sequence 37, Appl
33	822	82.9	166	4	US-09-552-265B-22	Sequence 22, Appl
34	822	82.9	166	4	US-09-552-265B-24	Sequence 24, Appl
35	820	82.7	165	4	US-09-604-871-1	Sequence 1, Appli
36	820	82.7	165	4	US-09-604-938-1	Sequence 1, Appli
37	818	82.5	166	4	US-09-552-265B-23	Sequence 23, Appl
38	818	82.5	166	4	US-09-552-265B-25	Sequence 25, Appl
39	817	82.4	166	4	US-09-552-265B-26	Sequence 26, Appl
40	816	82.3	166	4	US-09-552-265B-28	Sequence 28, Appl
41	813	82.0	166	4	US-09-552-265B-27	Sequence 27, Appl
42	812	81.9	166	4	US-09-552-265B-29	Sequence 29, Appl
43	810	81.7	166	4	US-09-552-265B-30	Sequence 30, Appl
44	809	81.6	166	4	US-09-552-265B-32	Sequence 32, Appl
45	806	81.3	166	4	US-09-552-265B-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-552-265B-34
; Sequence 34, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
; APPLICANT: DeSauvage, Frederick
; APPLICANT: Henner, Dennis, J.
; TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
; TITLE OF INVENTION: polypeptides and nucleic acids encoding the same
; FILE REFERENCE: GENENT.057CP1
; CURRENT APPLICATION NUMBER: US/09/552,265B
; CURRENT FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: US 09/307307
; PRIOR FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 82, 84
; OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-34

Query Match 99.6%; Score 987; DB 4; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.7e-113;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193
```

RESULT 2
US-09-552-265B-36
; Sequence 36, Application US/09552265B

```
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 82, 84
;   OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-36

  Query Match            99.3%;  Score 984;  DB 4;  Length 193;
  Best Local Similarity  99.5%;  Pred. No. 6.3e-113;
  Matches  192;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXTXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 3
US-09-552-265B-35
; Sequence 35, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 82, 84
;   OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-35

  Query Match            99.2%;  Score 983;  DB 4;  Length 193;
  Best Local Similarity  99.5%;  Pred. No. 8.4e-113;
  Matches  192;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXTXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAKKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 4
US-09-552-265B-37
; Sequence 37, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 37
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 82, 84
;   OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-37

  Query Match            98.9%;  Score 980;  DB 4;  Length 193;
  Best Local Similarity  99.0%;  Pred. No. 2e-112;
  Matches  191;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXTXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAKKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 5
US-07-903-220-1
; Sequence 1, Application US/07903220
; Patent No. 5322837
;  GENERAL INFORMATION:
;    APPLICANT:  Hewick, Rodney M.
;    TITLE OF INVENTION:  METHOD FOR THE PURIFICATION OF
;    TITLE OF INVENTION:  ERYTHROPOIETIN AND ERYTHROPOIETIN COMPOSITION
```

```
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Paul H. Heller
;      STREET:  Kenyon & Kenyon, One Broadway
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/903,220
;      FILING DATE:  19920731
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  1248/27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 429-1776
;      TELEFAX:  (202) 429-0796
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  193 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-07-903-220-1

  Query Match             98.2%;  Score 973;  DB 1;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 1.4e-111;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 6
US-08-883-795A-34
; Sequence 34, Application US/08883795A
; Patent No. 5985607
;  GENERAL INFORMATION:
;    APPLICANT:  Delcuve, Genevieve
;    APPLICANT:  Awang, Gregor
;    TITLE OF INVENTION:  Recombinant DNA Molecules and Expression
;    TITLE OF INVENTION:  Vectors for Tissue Plasminogen Activator
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BERESKIN & PARR
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,795A
;      FILING DATE:  27-JUN-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gravelle, Micheline
;      REGISTRATION NUMBER:  40,261
;      REFERENCE/DOCKET NUMBER:  7841-062
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  193 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-883-795A-34

  Query Match             98.2%;  Score 973;  DB 2;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 1.4e-111;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 7
US-09-552-265B-2
; Sequence 2, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
US-09-552-265B-2

  Query Match             98.2%;  Score 973;  DB 4;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 1.4e-111;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
               |||||||||||||||||||    |||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVRQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
               |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 8
US-09-552-265B-4
; Sequence 4, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-552-265B-4

  Query Match             98.2%;  Score 973;  DB 4;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 1.4e-111;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
               |||||||||||||||||||    |||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
               |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 9
US-09-552-265B-38
; Sequence 38, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 38
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 83, 85
;   OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-38

  Query Match             97.9%;  Score 970;  DB 4;  Length 193;
  Best Local Similarity   97.4%;  Pred. No. 3.3e-111;
  Matches  188;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
               |||||||||||||||||||    |||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMNXSXQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
               |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 10
US-09-552-265B-40
; Sequence 40, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 83, 85
;   OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-40

  Query Match             97.9%;  Score 970;  DB 4;  Length 193;
  Best Local Similarity   97.4%;  Pred. No. 3.3e-111;
  Matches  188;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
               |||||||||||||||||||    |||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMNXTXQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
```

```
Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 11
US-09-552-265B-5
; Sequence 5, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
US-09-552-265B-5

  Query Match             97.8%;  Score 969;  DB 4;  Length 193;
  Best Local Similarity   97.4%;  Pred. No. 4.4e-111;
  Matches  188;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVRQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAKKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 12
US-09-552-265B-39
; Sequence 39, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 83, 85
;   OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-39

  Query Match             97.5%;  Score 966;  DB 4;  Length 193;
  Best Local Similarity   96.9%;  Pred. No. 1e-110;
  Matches  187;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMNXSXQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAKKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 13
US-09-552-265B-41
; Sequence 41, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 83, 85
;   OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-41

  Query Match             97.5%;  Score 966;  DB 4;  Length 193;
  Best Local Similarity   96.9%;  Pred. No. 1e-110;
  Matches  187;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMNXTXQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAKKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 14
US-09-552-265B-42
; Sequence 42, Application US/09552265B
```

```
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 84, 86
;   OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-42

  Query Match             97.4%;  Score 965;  DB 4;  Length 193;
  Best Local Similarity   97.4%;  Pred. No. 1.4e-110;
  Matches  188;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||  :|  ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMENXSXAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 15
US-09-552-265B-44
; Sequence 44, Application US/09552265B
; Patent No. 6555343
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. 6555343el chimpanzee erythropoietin (chepo)
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP1
;  CURRENT APPLICATION NUMBER: US/09/552,265B
;  CURRENT FILING DATE:  2000-04-19
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 84, 86
;   OTHER INFORMATION: Xaa = any amino acid
US-09-552-265B-44

  Query Match             97.3%;  Score 964;  DB 4;  Length 193;
  Best Local Similarity   97.4%;  Pred. No. 1.8e-110;
  Matches  188;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||  :|  ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMENXTXAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


Search completed: September  8, 2003, 15:04:08
Job time : 13.2862 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 14:33:36 ; Search time 40 Seconds
(without alignments)
765.856 Million cell updates/sec

Title: US-09-813-775C-34
Perfect score: 991
Sequence: 1 MGVHECPAWLWLLLSLLSLP.....NFLRGKLKLYTGEACRTGDR 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	987	99.6	193	22	AAB35002	Chimpanzee erythro
2	987	99.6	193	23	AAE29323	Chimpanzee erythro
3	984	99.3	193	22	AAB35004	Chimpanzee erythro
4	984	99.3	193	23	AAE29325	Chimpanzee erythro
5	983	99.2	193	22	AAB35003	Chimpanzee erythro
6	983	99.2	193	23	AAE29324	Chimpanzee erythro
7	980	98.9	193	22	AAB35005	Chimpanzee erythro
8	980	98.9	193	23	AAE29326	Chimpanzee erythro
9	974	98.3	193	21	AAY94536	Human erythropoiet
10	973	98.2	193	6	AAP50300	Human erythropoiet
11	973	98.2	193	8	AAP70256	Sequence of human
12	973	98.2	193	15	AAR65499	Human prepro-eryth
13	973	98.2	193	16	AAR81982	Human erythropoiet
14	973	98.2	193	16	AAR71137	Human erythropoiet
15	973	98.2	193	16	AAR74141	Human erythropoiet
16	973	98.2	193	17	AAR98397	Human erythropoiet
17	973	98.2	193	21	AAY94530	Human erythropoiet
18	973	98.2	193	21	AAY93638	Amino acid sequenc
19	973	98.2	193	21	AAY99704	Human non-glycosyl
20	973	98.2	193	21	AAY43398	Human erythropoiet
21	973	98.2	193	22	AAB85573	Human erythropoiet
22	973	98.2	193	22	AAB34978	Human erythropoiet
23	973	98.2	193	23	AAE29299	Chimpanzee erythro
24	973	98.2	193	23	AAE15341	Human erythropoiet
25	973	98.2	193	24	AAE32131	Human erythropoiet
26	973	98.2	194	22	AAB34977	Chimpanzee erythro
27	973	98.2	330	13	AAR23076	Epo:IL-3 short, re
28	973	98.2	349	13	AAR23079	Epo:IL-3 Flex, rec
29	973	98.2	376	20	AAW99360	Human erythropoiet
30	970	97.9	193	7	AAP60598	Open reading frame
31	970	97.9	193	16	AAR74560	Erythropoietin. H
32	970	97.9	193	21	AAY94531	Human erythropoiet
33	970	97.9	193	22	AAB35006	Chimpanzee erythro
34	970	97.9	193	22	AAB35008	Chimpanzee erythro
35	970	97.9	193	23	AAE29327	Chimpanzee erythro
36	970	97.9	193	23	AAE29329	Chimpanzee erythro
37	969	97.8	193	16	AAR81986	Erythropoietin (Al
38	969	97.8	193	21	AAY94532	Human erythropoiet
39	969	97.8	193	21	AAY81913	Human EPO protein
40	969	97.8	193	23	AAE29300	Chimpanzee erythro
41	968	97.7	220	23	ABB79939	Human erythropoiet
42	967	97.6	193	18	AAW14143	Erythropoietin var
43	967	97.6	193	19	AAW58400	Human erythropoiet
44	967	97.6	193	22	AAB34979	Chimpanzee erythro
45	966	97.5	193	12	AAR11859	Erythropoietin ana

ALIGNMENTS

RESULT 1
AAB35002
ID AAB35002 standard; Protein; 193 AA.
XX
AC AAB35002;
XX
DT 27-MAR-2001 (first entry)
XX
DE Chimpanzee erythropoietin fragment SEQ ID NO: 34.
XX
KW Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
KW mapping; therapeutic agent.
XX
OS Pan sp.
XX
PN WO200068376-A1.
XX
PD 16-NOV-2000.
XX
PF 05-MAY-2000; 2000WO-US12370.
XX
PR 07-MAY-1999; 99US-0307307.
PR 28-MAR-2000; 2000US-0307307.
XX
PA (GETH) GENENTECH INC.
XX
PI Desauvage F, Henner DJ;
XX
DR WPI; 2001-007393/01.
XX
PT Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
PT of e.g. anemia, also derived proteins, antibodies and modulators -

```
XX
PS   Claim 35; Page 96; 109pp; English.
XX
CC   The present invention provides the coding and protein sequences of
CC   chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC   therapy, to block the activity of EPO, as hybridisation probes, in
CC   genetic and chromosome mapping and as therapeutic agents.
XX
SQ   Sequence   193 AA;

  Query Match             99.6%;  Score 987;  DB 22;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-99;
  Matches  193;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 2
AAE29323
ID   AAE29323 standard; Protein; 193 AA.
XX
AC   AAE29323;
XX
DT   27-JAN-2003  (first entry)
XX
DE   Chimpanzee erythropoietin (CHEPO) protein #3.
XX
KW   Chimpanzee; erythropoietin; gene therapy; blood disorder; immunoadhesin;
KW   chromosome identification; tissue typing; antianaemic; CHEPO.
XX
OS   Pan troglodytes.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 82
FT                   /label= Unknown
FT                   /note= "X is any amino acid except for proline"
FT   Misc-difference 84
FT                   /label= Unknown
FT                   /note= "X is any amino acid except for proline"
XX
PN   WO200274807-A2.
XX
PD   26-SEP-2002.
XX
PF   14-FEB-2002; 2002WO-US04773.
XX
PR   20-MAR-2001; 2001US-0813775.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Desauvage F,  Henner DJ;
XX
DR   WPI; 2002-759880/82.
XX
PT   New immunoadhesin comprising a chimpanzee erythropoietin (CHEPO)
PT   polypeptide, useful for the treatment of blood disorders with low or
PT   defective red blood cell production  -
XX
PS   Claim 2; Page 82; 120pp; English.
XX
CC   The invention relates to immunoadhesins comprising chimpanzee
CC   erythropoietin (CHEPO) polypeptide. The invention further relates to
CC   methods using and compositions comprising CHEPO immunoadhesins. The
CC   methods and compositions of the present invention are useful for the
CC   treatment of blood disorders characterised by low or defective red
CC   blood cell production. The CHEPO polypeptides may also be used as
CC   molecular weight markers, tissue typing and chromosome identification.
CC   CHEPO DNA is used in gene therapy. The present sequence is chimpanzee
CC   erythropoietin protein.
XX
SQ   Sequence   193 AA;

  Query Match             99.6%;  Score 987;  DB 23;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-99;
  Matches  193;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 3
AAB35004
ID   AAB35004 standard; Protein; 193 AA.
XX
AC   AAB35004;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Chimpanzee erythropoietin fragment SEQ ID NO: 36.
XX
KW   Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
KW   mapping; therapeutic agent.
XX
OS   Pan sp.
XX
PN   WO200068376-A1.
XX
PD   16-NOV-2000.
XX
PF   05-MAY-2000; 2000WO-US12370.
XX
PR   07-MAY-1999;   99US-0307307.
PR   28-MAR-2000; 2000US-0307307.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Desauvage F,  Henner DJ;
XX
DR   WPI; 2001-007393/01.
XX
PT   Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
PT   of e.g. anemia, also derived proteins, antibodies and modulators -
XX
PS   Claim 35; Page 96; 109pp; English.
XX
CC   The present invention provides the coding and protein sequences of
CC   chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC   therapy, to block the activity of EPO, as hybridisation probes, in
CC   genetic and chromosome mapping and as therapeutic agents.
```

```
XX
SQ   Sequence   193 AA;

  Query Match            99.3%;  Score 984;  DB 22;  Length 193;
  Best Local Similarity   99.5%;  Pred. No. 4.1e-99;
  Matches  192;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXTXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 4
AAE29325
ID   AAE29325 standard; Protein; 193 AA.
XX
AC   AAE29325;
XX
DT   27-JAN-2003  (first entry)
XX
DE   Chimpanzee erythropoietin (CHEPO) protein #5.
XX
KW   Chimpanzee; erythropoietin; gene therapy; blood disorder; immunoadhesin;
KW   chromosome identification; tissue typing; antianaemic; CHEPO.
XX
OS   Pan troglodytes.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 82
FT                   /label= Unknown
FT                   /note= "X is any amino acid except for proline"
FT   Misc-difference 84
FT                   /label= Unknown
FT                   /note= "X is any amino acid except for proline"
XX
PN   WO200274807-A2.
XX
PD   26-SEP-2002.
XX
PF   14-FEB-2002; 2002WO-US04773.
XX
PR   20-MAR-2001; 2001US-0813775.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Desauvage F,  Henner DJ;
XX
DR   WPI; 2002-759880/82.
XX
PT   New immunoadhesin comprising a chimpanzee erythropoietin (CHEPO)
PT   polypeptide, useful for the treatment of blood disorders with low or
PT   defective red blood cell production  -
XX
PS   Claim 2; Page 82; 120pp; English.
XX
CC   The invention relates to immunoadhesins comprising chimpanzee
CC   erythropoietin (CHEPO) polypeptide. The invention further relates to
CC   methods using and compositions comprising CHEPO immunoadhesins. The
CC   methods and compositions of the present invention are useful for the
CC   treatment of blood disorders characterised by low or defective red
CC   blood cell production. The CHEPO polypeptides may also be used as
CC   molecular weight markers, tissue typing and chromosome identification.
CC   CHEPO DNA is used in gene therapy. The present sequence is chimpanzee
CC   erythropoietin protein.
XX
SQ   Sequence   193 AA;

  Query Match            99.3%;  Score 984;  DB 23;  Length 193;
  Best Local Similarity   99.5%;  Pred. No. 4.1e-99;
  Matches  192;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXTXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 5
AAB35003
ID   AAB35003 standard; Protein; 193 AA.
XX
AC   AAB35003;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Chimpanzee erythropoietin fragment SEQ ID NO: 35.
XX
KW   Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
KW   mapping; therapeutic agent.
XX
OS   Pan sp.
XX
PN   WO200068376-A1.
XX
PD   16-NOV-2000.
XX
PF   05-MAY-2000; 2000WO-US12370.
XX
PR   07-MAY-1999;   99US-0307307.
PR   28-MAR-2000; 2000US-0307307.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Desauvage F,  Henner DJ;
XX
DR   WPI; 2001-007393/01.
XX
PT   Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
PT   of e.g. anemia, also derived proteins, antibodies and modulators -
XX
PS   Claim 35; Page 96; 109pp; English.
XX
CC   The present invention provides the coding and protein sequences of
CC   chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC   therapy, to block the activity of EPO, as hybridisation probes, in
CC   genetic and chromosome mapping and as therapeutic agents.
XX
SQ   Sequence   193 AA;

  Query Match            99.2%;  Score 983;  DB 22;  Length 193;
  Best Local Similarity   99.5%;  Pred. No. 5.3e-99;
  Matches  192;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAKKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193
```

RESULT 6
AAE29324
ID AAE29324 standard; Protein; 193 AA.
XX
AC AAE29324;
XX
DT 27-JAN-2003 (first entry)
XX
DE Chimpanzee erythropoietin (CHEPO) protein #4.
XX
KW Chimpanzee; erythropoietin; gene therapy; blood disorder; immunoadhesin;
KW chromosome identification; tissue typing; antianaemic; CHEPO.
XX
OS Pan troglodytes.
XX
FH Key Location/Qualifiers
FT Misc-difference 82
FT /label= Unknown
FT /note= "X is any amino acid except for proline"
FT Misc-difference 84
FT /label= Unknown
FT /note= "X is any amino acid except for proline"
XX
PN WO200274807-A2.
XX
PD 26-SEP-2002.
XX
PF 14-FEB-2002; 2002WO-US04773.
XX
PR 20-MAR-2001; 2001US-0813775.
XX
PA (GETH) GENENTECH INC.
XX
PI Desauvage F, Henner DJ;
XX
DR WPI; 2002-759880/82.
XX
PT New immunoadhesin comprising a chimpanzee erythropoietin (CHEPO)
PT polypeptide, useful for the treatment of blood disorders with low or
PT defective red blood cell production -
XX
PS Claim 2; Page 82; 120pp; English.
XX
CC The invention relates to immunoadhesins comprising chimpanzee
CC erythropoietin (CHEPO) polypeptide. The invention further relates to
CC methods using and compositions comprising CHEPO immunoadhesins. The
CC methods and compositions of the present invention are useful for the
CC treatment of blood disorders characterised by low or defective red
CC blood cell production. The CHEPO polypeptides may also be used as
CC molecular weight markers, tissue typing and chromosome identification.
CC CHEPO DNA is used in gene therapy. The present sequence is chimpanzee
CC erythropoietin protein.
XX
SQ Sequence 193 AA;

Query Match 99.2%; Score 983; DB 23; Length 193;
Best Local Similarity 99.5%; Pred. No. 5.3e-99;
Matches 192; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAKKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193
```

RESULT 7
AAB35005
ID AAB35005 standard; Protein; 193 AA.
XX
AC AAB35005;
XX
DT 27-MAR-2001 (first entry)
XX
DE Chimpanzee erythropoietin fragment SEQ ID NO: 37.
XX
KW Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
KW mapping; therapeutic agent.
XX
OS Pan sp.
XX
PN WO200068376-A1.
XX
PD 16-NOV-2000.
XX
PF 05-MAY-2000; 2000WO-US12370.
XX
PR 07-MAY-1999; 99US-0307307.
PR 28-MAR-2000; 2000US-0307307.
XX
PA (GETH) GENENTECH INC.
XX
PI Desauvage F, Henner DJ;
XX
DR WPI; 2001-007393/01.
XX
PT Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
PT of e.g. anemia, also derived proteins, antibodies and modulators -
XX
PS Claim 35; Page 96; 109pp; English.
XX
CC The present invention provides the coding and protein sequences of
CC chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC therapy, to block the activity of EPO, as hybridisation probes, in
CC genetic and chromosome mapping and as therapeutic agents.
XX
SQ Sequence 193 AA;

Query Match 98.9%; Score 980; DB 22; Length 193;
Best Local Similarity 99.0%; Pred. No. 1.1e-98;
Matches 191; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXTXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
```

```
Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAKKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193
```

RESULT 8
AAE29326
ID AAE29326 standard; Protein; 193 AA.
XX
AC AAE29326;
XX
DT 27-JAN-2003 (first entry)
XX
DE Chimpanzee erythropoietin (CHEPO) protein #6.
XX
KW Chimpanzee; erythropoietin; gene therapy; blood disorder; immunoadhesin;
KW chromosome identification; tissue typing; antianaemic; CHEPO.
XX
OS Pan troglodytes.
XX
FH Key Location/Qualifiers
FT Misc-difference 82
FT /label= Unknown
FT /note= "X is any amino acid except for proline"
FT Misc-difference 84
FT /label= Unknown
FT /note= "X is any amino acid except for proline"
XX
PN WO200274807-A2.
XX
PD 26-SEP-2002.
XX
PF 14-FEB-2002; 2002WO-US04773.
XX
PR 20-MAR-2001; 2001US-0813775.
XX
PA (GETH) GENENTECH INC.
XX
PI Desauvage F, Henner DJ;
XX
DR WPI; 2002-759880/82.
XX
PT New immunoadhesin comprising a chimpanzee erythropoietin (CHEPO)
PT polypeptide, useful for the treatment of blood disorders with low or
PT defective red blood cell production -
XX
PS Claim 2; Page 82; 120pp; English.
XX
CC The invention relates to immunoadhesins comprising chimpanzee
CC erythropoietin (CHEPO) polypeptide. The invention further relates to
CC methods using and compositions comprising CHEPO immunoadhesins. The
CC methods and compositions of the present invention are useful for the
CC treatment of blood disorders characterised by low or defective red
CC blood cell production. The CHEPO polypeptides may also be used as
CC molecular weight markers, tissue typing and chromosome identification.
CC CHEPO DNA is used in gene therapy. The present sequence is chimpanzee
CC erythropoietin protein.
XX
SQ Sequence 193 AA;

Query Match 98.9%; Score 980; DB 23; Length 193;
Best Local Similarity 99.0%; Pred. No. 1.1e-98;
Matches 191; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRNXTXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAKKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193
```

RESULT 9
AAY94536
ID AAY94536 standard; protein; 193 AA.
XX
AC AAY94536;
XX
DT 28-NOV-2000 (first entry)
XX
DE Human erythropoietin mutant E55N/G57T.
XX
KW Human; erythropoietin; Epo; glycosylation; anaemia;
KW chronic renal failure; myelosuppressive therapy; cancer;
KW viral infection; HIV; blood loss; mutein; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT /label= Signal
FT Misc-difference 55
FT /note= "Wild-type Glu substituted by Asn"
FT Modified-site 55..57
FT /note= "Asn is glycosylated"
FT Misc-difference 57
FT /note= "Wild-type Gly substituted by Thr"
XX
PN WO200024893-A2.
XX
PD 04-MAY-2000.
XX
PF 18-OCT-1999; 99WO-US24435.
XX
PR 23-OCT-1998; 98US-0178292.
XX
PA (AMGE-) AMGEN INC.
XX
PI Egrie JC, Elliott SG, Brown JK;
XX
DR WPI; 2000-350735/30.
XX
PT Raising and maintaining hematocrit in a mammal by administering an
PT effective amount of a hyperglycosylated analog of erythropoietin,
PT useful for treating anemia associated with myelosuppressive therapy or
PT excessive blood loss -
XX
PS Claim 27; Page -; 63pp; English.
XX
CC Human erythropoietin (Epo) is a glycoprotein hormone necessary
CC for the maturation of erythroid progenitor cells into erythrocytes.
CC It has been discovered that hyperglycosylated Epo has a longer
CC half-life and greater in vivo activity than recombinant human Epo.
CC The present sequence is the hyperglycosylated Epo mutant E55N/G57T.
CC This and several other hyperglycosylated Epo mutants (AAY94531 to
CC AAY94544) have been made by in vitro mutagenesis. Hyperglycosylated
CC Epo analogs are useful as they may be used instead of recombinant
CC Epo to increase and maintain the level of red blood cells in mammals.
CC The Epo analogs may be used to treat or prevent anaemia associated
CC with chronic renal failure, myelosuppressive therapy, certain cancers,
CC viral disease such as HIV and excessive blood loss.

```
CC   Note: The present sequence is not shown in the specification but is
CC   derived from the wild-type erythropoietin shown in Figure 1 (AAY94530).
XX
SQ   Sequence   193 AA;

  Query Match             98.3%;  Score 974;  DB 21;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 5.1e-98;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMNVTQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 10
AAP50300
ID   AAP50300 standard; protein; 193 AA.
XX
AC   AAP50300;
XX
DT   25-MAR-2003  (updated)
DT   01-JAN-1980  (first entry)
XX
DE   Human erythropoietin encoded by positve clone (phage lambda-hE1)
DE   isolated from human fetal liver gene bank.
XX
KW   Erythropoietin; red blood cell; erythrocyte; anaemia; blood;
KW   disorder; ss; phage lambda-hE1; gene bank.
XX
OS   Homo sapiens.
XX
PN   WO8502610-A.
XX
PD   20-JUN-1985.
XX
PF   11-DEC-1984;   84WO-US02021.
XX
PR   30-NOV-1984;   84US-0675298.
PR   13-DEC-1983;   83US-0561024.
PR   21-FEB-1984;   84US-0582185.
PR   28-SEP-1984;   84US-0655841.
XX
PA   (KIRI ) KIRIN AMGEN INC.
XX
DR   WPI; 1985-159229/26.
DR   N-PSDB; AAN50347.
XX
PT   New polypeptide having properties of erythropoietin - is prepd.
PT   by cultivation of transformed eucaryotic or procaryotic host
XX
PS   Disclosure; Page 43; 113pp; English.
XX
CC   Human erythropoietin encoded by a sequence encoded by this phage
CC   lambda-hE1 is essential for red blood cell formation and is used
CC   for the diagnosis and treatment of blood disorders such as anaemia.
CC   Large amounts of EPO may be obtained using recombinant DNA
CC   techniques in contrast to small amounts obtained from plasma
CC   and urine.  This sequence is expressed in E. coli.  See also
CC   AAN50345-6, AAN50348-50 and AAP50298-99, AAP50301.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   193 AA;

  Query Match             98.2%;  Score 973;  DB 6;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 6.5e-98;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 11
AAP70256
ID   AAP70256 standard; protein; 193 AA.
XX
AC   AAP70256;
XX
DT   19-FEB-1991  (first entry)
XX
DE   Sequence of human erythropoietin (EPO).
XX
KW   Renal anaemia therapy; hormone.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..27
FT                   /label=SIGNAL
FT   Protein         28..193
FT   Region          81..97
FT                   /note="Fragment that probe AAN70361 is based on"
XX
PN   EP232034-A.
XX
PD   12-AUG-1987.
XX
PF   19-JAN-1987;   87EP-0300399.
XX
PR   23-JAN-1986;   86JP-0012868.
XX
PA   (SUMO ) SUMITOMO CHEM IND KK.
PA   (SUMI-) SUMITOMI SEIYAKU KK.
XX
PI   Yanagi H,  Ogawa I,  Okamoto M,  Hozumi T,  Soga A,  Yoshima T;
PI   Tsutsumi M;
XX
DR   WPI; 1987-223006/32.
DR   N-PSDB; AAN70360, AAN70361.
XX
PT   Human erythropoietin prodn. - by culturing human cells, esp.
PT   Namalwa cells, transformed with DNA encoding human erythropoietin
XX
PS   Disclosure; Fig 1; 22pp; English.
XX
CC   A cDNA library was prepd. from the poly (A) RNA, which was isolated
CC   from the erythropoietin-producing human hepatoma cell Hp-1. The cDNA
CC   library was screened using the probes given in AAN70361 and AAN70362. A
CC   plasmid (named as p58-A20) was isolated. The nucleotide sequence of
CC   the cDNA obtained from this clone is shown in AAN70360.
XX
SQ   Sequence   193 AA;
```

```
  Query Match             98.2%;  Score 973;  DB 8;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 6.5e-98;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 12
AAR65499
ID   AAR65499 standard; Protein; 193 AA.
XX
AC   AAR65499;
XX
DT   25-MAR-2003  (updated)
DT   24-JUN-1995  (first entry)
XX
DE   Human prepro-erythropoietin.
XX
KW   Erythropoietin; therapeutic; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..27
FT                   /note= "leader peptide"
XX
PN   WO9425055-A1.
XX
PD   10-NOV-1994.
XX
PF   29-APR-1994;    94WO-US04755.
XX
PR   29-APR-1993;    93US-0055076.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   DeVries PJ,  Mellovitz BS,  Meuth JL,  Okasinski GF;
PI    Schaefer VG;
XX
DR   WPI; 1994-357906/44.
DR   N-PSDB; AAQ74760.
XX
PT   Erythropoietin analogues - useful for treatment of anaemia and
PT   have enhanced erythropoietic effect.
XX
PS   Disclosure; Page 38-39; 56pp; English.
XX
CC   DNA encoding human prepro-erythropoietin may be ligated into an
CC   expression vector for erythropoietin expression in a CHO
CC   cell culture.  Site-directed mutagenesis may be used in the
CC   construction of EPO analogues with improved activity, which may be
CC   used in pharmaceutical compositions for inducing erythropoiesis and
CC   treating anaemia.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   193 AA;

  Query Match             98.2%;  Score 973;  DB 15;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 6.5e-98;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 13
AAR81982
ID   AAR81982 standard; Protein; 193 AA.
XX
AC   AAR81982;
XX
DT   25-MAR-2003  (updated)
DT   27-FEB-1996  (first entry)
XX
DE   Human erythropoietin.
XX
KW   Erythropoietin; sialylation; sialic acid; glycosylation;
KW   reticulocyte; red blood cell; erythrocyte; haematocrit.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..27
FT                   /label= Sig_peptide
FT   Modified-site   51
FT                   /label= N-glycosylation_site
FT   Modified-site   65
FT                   /label= N-glycosylation_site
FT   Modified-site   110
FT                   /label= N-glycosylation_site
FT   Modified-site   153
FT                   /label= O-glycosylation_site
XX
PN   EP668351-A1.
XX
PD   23-AUG-1995.
XX
PF   12-OCT-1990;    95EP-0101849.
XX
PR   09-OCT-1990;    90WO-US05758.
PR   13-OCT-1989;    89US-0421444.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Byrne TE,  Elliott SG;
XX
DR   WPI; 1995-284791/38.
XX
PT   New human erythropoietin analogues with increased glycosylation -
PT   have increased activity useful for increasing prodn. of
PT   reticulocytes and red blood cells
XX
PS   Disclosure; Fig 5; 31pp; English.
XX
CC   Human urinary erythropoietin (AAR81982) is a glycoprotein contg. 3 N-
CC   linked and 1 O-linked oligosaccharide chain.  Erythropoietin
CC   analogues (AAR81983-87) have been produced in which the number of
CC   glycosylation sites is increased.
```

```
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   193 AA;

  Query Match             98.2%;  Score 973;  DB 16;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 6.5e-98;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 14
AAR71137
ID   AAR71137 standard; protein; 193 AA.
XX
AC   AAR71137;
XX
DT   25-MAR-2003  (updated)
DT   17-OCT-1995  (first entry)
XX
DE   Human erythropoietin.
XX
KW   Human erythropoietin; glycosylation; sialic acid; solubility;
KW   half-life; biological activity; proteolysis resistance; anaemia;
KW   chronic renal failure.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..27
FT                   /label= sig_peptide
XX
PN   WO9505465-A1.
XX
PD   23-FEB-1995.
XX
PF   16-AUG-1994;   94WO-US09257.
XX
PR   13-OCT-1989;   89US-0421444.
PR   17-AUG-1993;   93US-0108016.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Byrne TE,  Elliott SG;
XX
DR   WPI; 1995-098764/13.
XX
PT   Erythropoietin (EPO) analogues having additional glycosylation
PT   site(s) - to increase sialic acid content, thereby increasing
PT   solubility, serum half-life, biological activity and resistance
PT   to proteolysis.
XX
PS   Disclosure; Pages 80-81; 108pp; English.
XX
CC   AAR71137 describes the amino acid sequence of human erythropoietin
CC   (EPO), from which the inventions novel human EPO analogues were
CC   derived. The analogues have at least one additional glycosylation
CC   site, this is used to increase the sialic acid content which in
CC   turn increases the solubility, half-life, biological activity and
CC   proteolysis resistance of the protein. The analogues are useful
CC   in claimed compsns. for the treatment of chronic renal failure
CC   associated anaemia.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   193 AA;

  Query Match             98.2%;  Score 973;  DB 16;  Length 193;
  Best Local Similarity   97.9%;  Pred. No. 6.5e-98;
  Matches  189;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193


RESULT 15
AAR74141
ID   AAR74141 standard; Protein; 193 AA.
XX
AC   AAR74141;
XX
DT   25-MAR-2003  (updated)
DT   30-OCT-1995  (first entry)
XX
DE   Human erythropoietin.
XX
KW   Erythropoietin; anemia; gene therapy; gene transfer; red blood cell;
KW   RBC; erythrocyte; transformation; myoblast; EPO.
XX
OS   Homo sapiens.
XX
PN   WO9513376-A1.
XX
PD   18-MAY-1995.
XX
PF   09-NOV-1994;   94WO-US13066.
XX
PR   07-OCT-1994;   94US-0320480.
PR   10-NOV-1993;   93US-0149871.
XX
PA   (AMGE-) AMGEN INC.
PA   (UYSC-) UNIV SOUTHERN CALIFORNIA.
XX
PI   Hamamori Y,  Kedes LH,  Samal BB;
XX
DR   WPI; 1995-194095/25.
DR   N-PSDB; AAQ92296.
XX
PT   Gene therapy for treatment of anaemia  - and increasing red blood cell
PT   production by transforming red blood cells with the erythropoietin gene
XX
PS   Disclosure; Page 38-40; 51pp; English.
XX
CC   The amino acid sequence encoded by human EPO cDNA is given in AAR74141.
CC   Transfection of target cells, e.g. myoblasts, with EPO cDNA and
CC   implantation into muscle tissue provides increased RBC prodn.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   193 AA;
```

Query Match 98.2%; Score 973; DB 16; Length 193;
Best Local Similarity 97.9%; Pred. No. 6.5e-98;
Matches 189; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC 60

Qy         61 SLNENITVPDTKVNFYAWKRNXSXQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120
              |||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db         61 SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL 120

Qy        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 180

Qy        181 KLYTGEACRTGDR 193
              |||||||||||||
Db        181 KLYTGEACRTGDR 193
```

Search completed: September 8, 2003, 15:00:47
Job time : 41 secs

	Type	L #	Hits	Search Text	DBs	Time Stamp	Comments	Error Definition	Errors
1	BRS	L1	8451	erythropoietin	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/09/08 15:36			0
2	BRS	L2	6	erythropoietin same chimpanzee	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/09/08 15:44			0
3	BRS	L3	4	(chimeric or fusion) same 2	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/09/08 15:45			0
4	BRS	L4	4768	epitope adj tag	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/09/08 15:45			0
5	BRS	L5	6843	immunoglobulin same fc	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/09/08 15:46			0
6	BRS	L6	12746	immunoglobulin same (igg or igg1)	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/09/08 15:46			0
7	BRS	L7	0	3 same (4 or 5 or 6)	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/09/08 15:46			0

```
FILE 'MEDLINE' ENTERED AT 15:51:17 ON 08 SEP 2003

FILE 'CAPLUS' ENTERED AT 15:51:17 ON 08 SEP 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'BIOSIS' ENTERED AT 15:51:17 ON 08 SEP 2003
COPYRIGHT (C) 2003 BIOLOGICAL ABSTRACTS INC.(R)

FILE 'EMBASE' ENTERED AT 15:51:17 ON 08 SEP 2003
COPYRIGHT (C) 2003 Elsevier Science B.V.  All rights reserved.

FILE 'SCISEARCH' ENTERED AT 15:51:17 ON 08 SEP 2003
COPYRIGHT 2003 THOMSON ISI

FILE 'AGRICOLA' ENTERED AT 15:51:17 ON 08 SEP 2003

=> s erythropoietin (p) chimpanzee
L1            7 ERYTHROPOIETIN (P) CHIMPANZEE

=> s L1 (p) (chimeric or fusion)
L2            4 L1 (P) (CHIMERIC OR FUSION)

=> duplicate remove l2
DUPLICATE PREFERENCE IS 'CAPLUS, BIOSIS'
KEEP DUPLICATES FROM MORE THAN ONE FILE?  Y/(N):n
PROCESSING COMPLETED FOR L2
L3             4 DUPLICATE REMOVE L2 (0 DUPLICATES REMOVED)

=> d l3 1-4 ibib abs

L3   ANSWER 1 OF 4  CAPLUS  COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER:         2003:222255  CAPLUS
DOCUMENT NUMBER:          138:248956
TITLE:                    Methods for the modulation of erythropoiesis with
                          chimpanzee erythropoietin (CHEPO) polypeptides and
                          nucleic acids
INVENTOR(S):              Desauvage, Frederic; Henner, Dennis J.
PATENT ASSIGNEE(S):       USA
SOURCE:                   U.S. Pat. Appl. Publ., 81 pp., Cont.-in-part of U.S.
                          Ser. No. 552,265.
                          CODEN: USXXCO
DOCUMENT TYPE:            Patent
LANGUAGE:                 English
FAMILY ACC. NUM. COUNT:   3
PATENT INFORMATION:

      PATENT NO.        KIND  DATE           APPLICATION NO.  DATE
      ----------------  ----  --------       ---------------  --------
      US 2003054494      A1   20030320       US 2001-813775   20010320
      US 6555343         B1   20030429       US 2000-552265   20000419
      WO 2002074807      A2   20020926       WO 2002-US4773   20020214
      WO 2002074807      A3   20030724
          W:  AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,
              CO, CR, CU, CZ, CZ, DE, DE, DK, DK, DM, DZ, EC, EE, EE, ES, FI,
              FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP,
              KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
              MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK, SL,
              TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, AM, AZ,
              BY, KG, KZ, MD
          RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH,
              CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR,
              BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG
PRIORITY APPLN. INFO.:                    US 2000-552265   A2 20000419
                                          US 2000-287594P  P  20000328
                                          US 2001-813775   A  20010320
AB    The present invention is directed to novel   ***chimpanzee***
      ***erythropoietin***   polypeptides and to nucleic acid mols. encoding
      those polypeptides.  Also provided herein are vectors and host cells
      comprising those nucleic acid sequences,   ***chimeric***   polypeptide
      mols. comprising the polypeptides of the present invention fused to
      heterologous polypeptide sequences, and antibodies which bind to the
      polypeptides of the present invention.
```

L3 ANSWER 2 OF 4 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. on STN
ACCESSION NUMBER: 2003:248663 BIOSIS
DOCUMENT NUMBER: PREV200300248663
TITLE: Chimpanzee erythropoietin (CHEPO) polypeptides and nucleic acids encoding the same.
AUTHOR(S): DeSauvage, Frederic (1); Henner, Dennis J.
CORPORATE SOURCE: (1) Foster City, CA, USA USA
ASSIGNEE: Genentech Inc.
PATENT INFORMATION: US 6555343 April 29, 2003
SOURCE: Official Gazette of the United States Patent and Trademark Office Patents, (Apr. 29 2003) Vol. 1269, No. 5, pp. No Pagination. http://www.uspto.gov/web/menu/patdata.html. e-file.
ISSN: 0098-1133.
DOCUMENT TYPE: Patent
LANGUAGE: English
AB The present invention is directed to novel ***chimpanzee*** ***erythropoietin*** polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, ***chimeric*** polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

L3 ANSWER 3 OF 4 CAPLUS COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER: 2002:736286 CAPLUS
DOCUMENT NUMBER: 137:257948
TITLE: Chimpanzee erythropoietin (CHEPO) - immunoadhesins for use in regulating erythropoiesis
INVENTOR(S): Desauvage, Frederic; Henner, Dennis J.
PATENT ASSIGNEE(S): Genentech, Inc., USA
SOURCE: PCT Int. Appl., 120 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 3
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002074807	A2	20020926	WO 2002-US4773	20020214
WO 2002074807	A3	20030724		
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, CZ, DE, DE, DK, DK, DM, DZ, EC, EE, EE, ES, FI, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD			
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				
US 2003054494	A1	20030320	US 2001-813775	20010320

PRIORITY APPLN. INFO.: US 2001-813775 A 20010320
US 2000-552265 A2 20000419

AB The present invention is directed to immunoadhesins comprising chimpanzee erythropoietin (CHEPO) polypeptides. The immunoadhesins have an enhanced in vivo half-life compared to the corresponding CHEPO polypeptide and retains CHEPO biol. activity. Also provided herein are nucleic acid mols. encoding such immunoadhesins, vectors and host cells comprising those nucleic acid sequences, and methods using and compns. comprising the CHEPO immunoadhesins.

L3 ANSWER 4 OF 4 CAPLUS COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER: 2000:814616 CAPLUS
DOCUMENT NUMBER: 134:1044
TITLE: Chimpanzee erythropoietin and cDNA and methods of modulating erythropoiesis
INVENTOR(S): Desauvage, Frederic; Henner, Dennis J.
PATENT ASSIGNEE(S): Genentech, Inc., USA
SOURCE: PCT Int. Appl., 109 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 3
PATENT INFORMATION:

```
     PATENT NO.          KIND  DATE        APPLICATION NO.   DATE
     --------------      ----  --------    --------------    --------
     WO 2000068376        A1   20001116      WO 2000-US12370  20000505
         W:  AE, AG, AL, AM, AT, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN,
             CR, CU, CZ, CZ, DE, DE, DK, DK, DM, DZ, EE, EE, ES, FI, FI, GB,
             GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KR,
             KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO,
             NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK, SL, TJ, TM, TR, TT,
             TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU,
             TJ, TM
         RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE,
             DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF,
             CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
     US 6555343           B1   20030429      US 2000-552265   20000419
     EP 1177285           A1   20020206      EP 2000-928879   20000505
         R:  AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,
             IE, SI, LT, LV, FI, RO
     JP 2002543784        T2   20021224      JP 2000-616342   20000505
PRIORITY APPLN. INFO.:                    US 1999-307307   A  19990507
                                          US 2000-287594P  P  20000328
                                          US 2000-552265   A  20000419
                                          WO 2000-US12370  W  20000505
AB   The present invention is directed to novel   ***chimpanzee***
       ***erythropoietins***   and to nucleic acid mols. encoding those
     polypeptides.  Also provided herein are vectors and host cells comprising
     those nucleic acid sequences,   ***chimeric***   proteins comprising the
       ***erythropoietins***   of the present invention fused to heterologous
     protein sequences, antibodies which bind to the proteins of the invention,
     and to methods for producing the proteins of the invention.  Agonists and
     antagonists of   ***erythropoietin***   may be used to modulate
     erythropoiesis.
REFERENCE COUNT:          8     THERE ARE 8 CITED REFERENCES AVAILABLE FOR THIS
                                RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT


=> s epitope tag
L4         1990 EPITOPE TAG

=> s immunoglobulin (p) fa
L5         1133 IMMUNOGLOBULIN (P) FA

=> s immunoglobulin (p) igg or igg1
L6       124027 IMMUNOGLOBULIN (P) IGG OR IGG1

=> s immunoglobulin (p) (igg or igg1)
L7        93467 IMMUNOGLOBULIN (P) (IGG OR IGG1)

=> s immunoglobulin (p) fc
L8        27914 IMMUNOGLOBULIN (P) FC

=> s l4 and (l7 otr l8)
MISSING OPERATOR L7 OTR
The search profile that was entered contains terms or
nested terms that are not separated by a logical operator.

=> s l4 and (l7 or l8)
L9           34 L4 AND (L7 OR L8)

=> d his

     (FILE 'HOME' ENTERED AT 15:50:53 ON 08 SEP 2003)

     FILE 'MEDLINE, CAPLUS, BIOSIS, EMBASE, SCISEARCH, AGRICOLA' ENTERED AT
     15:51:17 ON 08 SEP 2003
L1             7 S ERYTHROPOIETIN (P) CHIMPANZEE
L2             4 S L1 (P) (CHIMERIC OR FUSION)
L3             4 DUPLICATE REMOVE L2 (0 DUPLICATES REMOVED)
L4          1990 S EPITOPE TAG
L5          1133 S IMMUNOGLOBULIN (P) FA
L6        124027 S IMMUNOGLOBULIN (P) IGG OR IGG1
L7         93467 S IMMUNOGLOBULIN (P) (IGG OR IGG1)
L8         27914 S IMMUNOGLOBULIN (P) FC
L9            34 S L4 AND (L7 OR L8)

=> s L3 and (l4 or l7 or l8)
L10            3 L3 AND (L4 OR L7 OR L8)
```

```
=> s desauvage frederic/au
L11          4 DESAUVAGE FREDERIC/AU

=> s henner dennis/au
L12         12 HENNER DENNIS/AU

=> s l11 or l12
L13         16 L11 OR L12

=> s l13 and l3
L14          4 L13 AND L3

=> s l14 not l3
L15          0 L14 NOT L3

=> d his

     (FILE 'HOME' ENTERED AT 15:50:53 ON 08 SEP 2003)

     FILE 'MEDLINE, CAPLUS, BIOSIS, EMBASE, SCISEARCH, AGRICOLA' ENTERED AT
     15:51:17 ON 08 SEP 2003
L1            7 S ERYTHROPOIETIN (P) CHIMPANZEE
L2            4 S L1 (P) (CHIMERIC OR FUSION)
L3            4 DUPLICATE REMOVE L2 (0 DUPLICATES REMOVED)
L4         1990 S EPITOPE TAG
L5         1133 S IMMUNOGLOBULIN (P) FA
L6       124027 S IMMUNOGLOBULIN (P) IGG OR IGG1
L7        93467 S IMMUNOGLOBULIN (P) (IGG OR IGG1)
L8        27914 S IMMUNOGLOBULIN (P) FC
L9           34 S L4 AND (L7 OR L8)
L10           3 S L3 AND (L4 OR L7 OR L8)
L11           4 S DESAUVAGE FREDERIC/AU
L12          12 S HENNER DENNIS/AU
L13          16 S L11 OR L12
L14           4 S L13 AND L3
L15           0 S L14 NOT L3

=> log y
COST IN U.S. DOLLARS                            SINCE FILE      TOTAL
                                                   ENTRY    SESSION
FULL ESTIMATED COST                                50.04      50.25

DISCOUNT AMOUNTS (FOR QUALIFYING ACCOUNTS)      SINCE FILE      TOTAL
                                                   ENTRY    SESSION
CA SUBSCRIBER PRICE                                -1.95      -1.95

STN INTERNATIONAL LOGOFF AT 15:58:48 ON 08 SEP 2003
```